Wavelet-Based Image Registration and Segmentation Framework for the Quantitative Evaluation of Hydrocephalus

By
Fan Luo

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the Degree of Master of Science in Applied Science

Dec, 2006, Halifax, Nova Scotia

Your file Votre référence
ISBN: 978-0-494-23854-7
Our file Notre référence
ISBN: 978-0-494-23854-7

Certification

Name: Fan Luo

Degree: Master of Science in Applied Science

Title of Thesis: Wavelet-Based Registration and Segmentation Framework for the Quantitative Evaluation of Hydrocephalus

Examining Committee:

Dr. Qigang Gao, External Examiner
Dalhousie University

Dr. Norma Linney, Senior Supervisor (Math & Computing Science)

Dr. Sageev Oore, Supervisory Committee (Math & Computing Science)

Dr. Peter Gregson, Supervisory Committee
(Electrical and Computer Engineering, Dalhousie)

Dr. Matthias Schmidt, Collaborator (IWK Health Centre)

Dr. Malcolm Butler, Program Co-ordinator

Dr. Kevin Vessey, Dean of Graduate Studies

Date Certified: December 13, 2006

Wavelet-Based Image Registration and Segmentation Framework for the Quantitative Evaluation of Hydrocephalus

by Fan Luo

Abstract

Hydrocephalus, a condition of increased fluid in the brain, is traditionally diagnosed by a visual assessment of CT scans. This thesis developed a quantitative measure of the change in ventricular volume over time. The framework includes: adaptive registration based on mutual information and wavelet multiresolution analysis, adaptive segmentation with a novel feature extraction method based on Dual-Tree Complex Wavelet Transform (DT-CWT) coefficients, and a volume calculation. The framework, when tested on physical phantoms had volume calculation accuracy of 1.0%. When tested on 8 clinical cases, the results reflected and predicted the diagnosis of the doctors, with less than 5% calculated volume change for cases where the diagnosis indicated the patient was stable, and more than 20% calculated volume change for cases for which hydrocephalus had been diagnosed. The outcome illustrated that the framework has good potential for development as a tool to aid in the diagnosis of hydrocephalus.

Dec 19, 2006

ACKNOWLEDGEMENTS

In order to finish this thesis, I received all kinds of help from many persons. I want to express the sincere appreciation to them.

First of all, I would like to thank Dr. Norma Linney who is my supervisor during my graduate study. She provided me not only the helpful ideas and comments but also the correct way to do research. She made me understand what real research is.

Secondly, I want to thank Dr. Schmidt who provided all the clinic image data for the testing procedure. Also, I will give my thanks to my committee members: Dr. Sageev Oore and Dr. Peter Gregson because of their helpful comments on my thesis.

Third, thank you to all my friends and professors from the Department of Mathematics and Computing Science, Saint Mary's University. I had a beautiful memory here. Special thanks to Dr. Pawan Lingras for giving me the chance to be enrolled in this Master's program and Owen Smith for the excellent technical support. Moreover, I would like to acknowledge to the financial support from the Graduate Studies Office and the Natural Sciences and Engineering Research Council for their support.

Finally, I thank my parents for the support during these two years.

Contents

List of Tables

List of Figures

List of Symbols

A_i: The area of the ith slice

α: A parameter vector for the affine transform

α^*: A optimal parameter vector for the affine transform

α_p: The perfect transformation parameter vector

$\alpha(\sigma)$: The angle between two gradient vectors

$CWT_x^{\psi}(\tau, s)$: Continuous Wavelet Transform

CWT: Complex wavelet transform

$D_{i,\theta}(x, y)$: The subband for $DT - CWT$

$DT - CWT$: Dual-Tree Complex Wavelet Transform

D_{volume}: Difference between two volumes

$D_{ref-float}$: The difference between the reference image and the floating image

$D_{ref-reg}$: The difference between the reference image and the registered image

E_{tex}: Texture energy

$\epsilon_B(f)$: Morphological erosion

$f(x, y)$: An image

$f_{Activity}$: Texture activity measure

$f_{bilinear}$: Function of bilinear interpolation

f_{interp}: Function of interpolation

$f_{MediFilt_\theta}$: Directional median filtering along θ

f_{weight}: Weight function

$\bigtriangledown f$: Magnitude of $\bigtriangledown F$

$\bigtriangledown F$: Gradient information of an image

$F_{float}(x, y)$: Floating image in registration

$F_{ref}(x, y)$: Reference image in registration

$F_{reg}(x, y)$: Registered image in registration

$G(A, B)$: The gradient term for registration

G_x: Gradient along x axis

G_y: Gradient along y axis

$G_{watershed}(x, y)$: The modified gradient information for the watershed transform

$H(A)$: Shannon entropy of image A

$H(A, B)$: Joint entropy between images A and B

$H(\mu, v)$: Gaussian function in the frequency domain

$I(A, B)$: Variable representing mutual information between images A and B

$I_{new}(A, B)$: Modified mutual information function that includes gradient information

IR: The improvement rate for registration

$M_{Erosion}$: Morphological erosion operation

$M_{Dilation}$: Morphological dilation operation

MI: Acronym for mutual information

$NMI(A, B)$: Normalized mutual information

$NMI_{new}(A, B)$: Modified normalized mutual information function that includes gradient information

p_i: Probability of a given symbol

$\psi_{j,k}(t)$: Wavelets

ψ: "Mother" wavelet

$\bigtriangledown\rho$: The phase component of gradient information

$R_{half}(\zeta)$: Half-wave rectification for segmentation

S: Similarity index

S_1, S_2: Set 1 and Set 2

S_{avg}: The average similarity index

$S_{thickness}$: Slice thickness

S_{ks}: Similarity matrix based on the Kolmogorov-Smirnov Test

$S_{i,\theta}(x, y)$: 2D directional median filtering results

$G^{tex}_{i,\theta}(x, y)$: The modified gradient information of the DT-CWT oriented at the θ at ith scale level

$G^{tex}(x, y)$: The texture gradient information

t_1: First scanning at time

t_2: Second scanning at time

TP: True positive

TN: True negative

FP: False positive

FN: False negative

V_i: The volume of the ith slice

V: The volume of the ventricles

v_{t_1}: Volume at time t_1

v_{t_2}: Volume at time t_2

$\omega_{i,\theta}$: The weighting function for calculating texture gradient

WT: Wavelet Transform

Chapter 1

Introduction

Hydrocephalus means "water brain". It is the result of excessive accumulation of fluid around the brain. Traditionally, hydrocephalus has been described as a condition characterized by increased intracranial pressure (ICP), increased cerebrospinal fluid (CSF) volume, and dilatation of the CSF spaces known as cerebral ventricles [35]. This condition usually appears in childhood, however, the condition can occur as a result of obstruction to normal CSF flow at any age. The clinical approach consists of early identification of the condition, followed by shunting of CSF to relieve intracranial pressure.

The prominent feature of hydrocephalus is the significant volume change of the ventricular system. This volume change can be seen in radiological images. Traditionally, estimation of the size of the ventricular system is done visually by doctors. But visual assessment has limited accuracy since the shape of the ventricular system is very complex. Moreover, differences in the orientations of slices from one study

to the next makes direct visual comparison of serial imaging studies difficult. This problem is refered to as the displacement problem. In this thesis, computer vision techniques were used to calculate the change in the volume of the ventricles with the goal of providing doctors with a tool to aid in the diagnosis of hydrocephalus.

There are four main techniques discussed in the thesis: wavelet-based analysis, image registration, image segmentation, and volume calculation. Considerable research has been done in these four areas. Wavelet multiresolution decomposition was used to make the registration process easier at different scales. Wavelet coefficients were applied to extract different features from the different textures. Mutual information-based registration was introduced to solve the displacement problem between two scans that were taken at different times. Morphological watershed segmentation was modified to automatically segment the CT images into several different homogenous regions. A volume calculation algorithm was applied to compute the ventricular volume using a stack of 2D images.

This thesis is organized in the following manner. In Chapter 2, a brief literature review is given on the following topics: wavelet applications in medical image processing, mutual information registration, image segmentation methods, and volume calculation algorithms. Chapter 3 gives the background theory on which the project and algorithms are based. Chapter 4 is a detailed description of the experimental methods. Experimental results are given in Chapter 5 with both data tables and some result images. In Chapter 6, some conclusions based on the experimental results are drawn. Future work is introduced in the last chapter.

Chapter 2

Background

Generally, four main kinds of computer vision techniques were used in the thesis: wavelet multiresolution analysis, image registration, image segmentation, and volume calculation. There has been much research work done in these fields. The following sections will briefly review the literature.

2.1 Wavelet Application

From the 1980's to the present, wavelet techniques have grown rapidly and found many applications in medical image processing. Because of its multiresolution characteristics, wavelet decomposition is very helpful for analyzing small areas (details) in an image. Wavelets have many applications in medical image processing. Basically, they can be applied to denoising [18][20][33], enhancement [18][45][8], and feature extraction [1][34]. In this thesis, feature extraction based on wavelet coefficients and wavelet multiresolution analysis is a key component of the registration and segmen-

tation algorithms.

Many researchers have made significant progress using wavelets for feature extraction. Saxena *et al.* [34] used two kinds of wavelets to extract small features from electrocardiogram (ECG) signals. One is the quadratic spline wavelet (QSWT) and the other is Daubechies six-coefficient (DU6) wavelet. The use of a QSWT provides a fixed correlation between the ECG characteristic points and the modulus maxima lines. DU6 has smooth features, so it is more suitable to detecting wave fiducials.

Common wavelet transforms can represent an image in three different directions, horizontal, vertical, and diagonal. For some applications, three selective directions are good enough for extracting features from an image or a texture. However, for some applications, such as texture classification and image retrieval, the common wavelet transform does not work very well. Furthermore, the common wavelet transform causes aliasing and exhibits shift variance [23]. As a result, many modified wavelet transforms have been introduced. The most famous examples are the complex wavelet transform, the tree-structured wavelet transform, and the dual-tree complex wavelet transform (DT-CWT). Kingsbury [23] introduced the complex wavelet transform to effectively solve the two problems of aliasing and shift variance. The complex wavelet transform filters have complex coefficients and generate complex output samples. Since the output sample rates are unchanged from the discrete wavelet transform, but each sample contains a real and imaginary part, a redundancy of $2:1$ is introduced in one dimension and 4:1 in two dimensions. The results show that the complex wavelet transform is approximately shift invariant and has better selective directions than the

dicrete wavelet transform. For example, a 2-D complex wavelet transform produces six bandpass subimages of complex coefficients at each level that are strongly oriented at angles of $\pm 15°$, $\pm 45°$, and $\pm 75°$.

For many applications, it is important that the transform be perfectly invertible. Unfortunately, it is very difficult to design an inverse transform based on complex filters. Moreover, the inefficient computation and redundancy are two disadvantages of the complex wavelet transform. Therefore, Kingsbury [24] designed the dual-tree complex wavelet transform (DT-CWT) which added perfect reconstruction, limited redundancy and efficient order-N computation to the other attractive properties of complex wavelets: shift invariance and good directional selectivity. Similar to the 2-D complex wavelet transform, the 2-D DT-CWT is approximately shift invariant and has six bandpass subimages at each level. The DT-CWT has been shown to be suitable for feature extraction. Hill *et al.* [16] and O'Callaghan *et al.* [30] used DT-CWT coefficients to construct the texture gradient map for the watershed segmentation algorithm. Their results showed that the DT-CWT coefficients represent the texture's features well and improve the results of the watershed method.

2.2 Registration

Image registration is the process of estimating an optimal transformation between two images. In this thesis, registration is used to compensate for the displacement of the human head from one CT scan to a second scan of the same patient taken at a different time. The displacement problem creates difficulty when trying to compare

the two sets of resulting images since the human head will not be in exactly the same position at the two scanning times. Generally, registration can be divided into two categories: rigid-body registration and non-rigid body registration. Rigid-body registration uses an affine transformation. Non-rigid body registration is based on a deformable model transformation. In this thesis, since the skull in the images is a rigid body and we are interested in how it has turned from one set of images to the next, only rigid body registration was considered. A deformable model transformation would distort the results.

Many studies have been performed on image registration. One of the approaches used is feature-based registration, in which significant regions, lines, or points are chosen as the landmarks. In Sun's work [39], the eyeballs were picked as the landmarks in the two images. Then the size of the two eyeballs and the distance between the nose and eyeball centers were measured. A geometric method was then used to compute the angle between these two pairs of eyeballs. Feature-based registration has the disadvantages of being linear and dependant on the user's manual choice of landmarks. Because of these disadvantages, mutual information (MI) registration was introduced both by Collignon [7] and by Viola and Wells [44].

Mutual information measures the mutual dependence of one imge on another or the information in image 1 that is shared by image 2. The larger the mutual information, the more two images look alike. Therefore, the global maximum of the mutual information function is used as the criterion for optimum registration. Since MI is based on Shannon Entropy estimation, it is an automatic, intensity-based

measure which does not need any landmarks or features defined. It does not assume a linear relationship among the gray values in the images. Research has shown that mutual information is suitable for medical image registration [27] and displacement correction [22].

Sometimes, mutual information used on its own, results in misregistration [38] because of the influence of the local maxima. So, normalized mutual information (NMI) [38] and entropy correlation coefficients (ECC) [25] are proposed by researchers. According to the outcome, NMI and ECC can reduce the influence of local maximum somewhat. However, it is still not easy to find the global maxima in the MI function. Therefore, some researchers incorporate the dependence of the gray values on neighboring voxels, in other words, the spatial information, and in particular, gradient spatial information. This method improves the registration results significantly. Pluim *et al.* [32] proposed a modified mutual information function which considers the gradient information of the image. This method multiplies the mutual information with a gradient term that is based on both the magnitude and orientation of the gradients. The results show that the MI function is much smoother than before gradients were incorporated. Some multiresolution methods, such as spline pyramids [41] and the 2-level multiresolution hierarchy [26], are also suggested to improve the performance of mutual information.

The registration measure defines an n-dimensional functional transformation with n degrees of freedom. The optimization of this function is assumed to correspond to the transformation that correctly registers the images. Two optimization methods are

popularly used in image registration: the simplex method and Powell's routine. If the relationship is linear, Powell's routine [11], which can optimize each transformation parameter in turn and does not require function derivatives to be calculated, can perform well with an acceptable optimization time. However, it is sensitive to local optima in the registration function. On the other hand, if the relationship is non-linear, it is better to use the simplex method [28] which considers all degrees of freedom simultaneously. The simplex method is not known for its speed of convergence.

In this thesis, the displacement problem in CT images might be caused by more than one parameter at a time. In other words, all the parameters used for optimization might be nonlinear for the registration function. In order to handle this probable situation, the simplex method was chosen, even though it is very time-consuming.

2.3 Segmentation

Image segmentation is the process of separating out mutually exclusive homogeneous regions of interest. For example, in this thesis, the images include mainly four parts: the background, the skull, the soft tissues inside the skull, and the ventricles. Different parts have different textures and the texture inside each part is considered to be homogeneous. The goal of the segmentation algorithm is to distinguish the different textures. There are many segmentation techniques, for example, histogram thresholding, region-growing, morphological watersheds, and techniques based on statistical approaches and wavelet coefficients.

2.3.1 Thresholding

Thresholding is the simplest segmentation algorithm and is implemented by selecting appropriate gray levels as thresholds. If a pixel lies above the threshold it will be marked as foreground, otherwise, it is marked as background. Usually, the threshold is selected by analyzing the histogram of the image. Cheriet *et al.* [6] presented a general recursive approach to thresholding which is based on discriminant analysis for image segmentation by extending Otsu's method [31]. Cheriet's approach segments the brightest homogeneous object from a given image at each recursion, leaving the darkest homogeneous object. The thresholding operation is regarded as the partitioning of pixels of an image into two classes: object and background. The conclusion shows that the method gives good results when the target object is the darkest object in a given image. However, when the target object is not darkest, the method fails to segment properly.

2.3.2 Region growing

Region growing is another popular segmentation algorithm [39]. It takes one or more seed pixels, and grows the regions around them based upon a certain homogeneity criterion. If the adjoining pixels are similar to the seeds, the pixels are merged within a single region. The process continues until all the pixels in the image are assigned to one or more regions. For a project similar to ours, Sun [39] used region growing with two criteria to segment the ventricles. Her method starts with a seed pixel which belongs to the ventricle, and then iteratively checks to decide if each neighboring pixel

belongs to the ventricle or not. For a pixel to be included in the region, the absolute gray-level difference between the pixel and the seed has to be less than a threshold, and the pixel has to be 8-connected to at least one pixel in that region.

Chang and Li [5] proposed a region-growing framework for image segmentation. This process is guided by regional feature analysis and no parameter tuning or *a priori* knowledge about the image is required. The algorithm is known as the Fast Adaptive Segmentation (FAS) algorithm. The focus of this study is on investigating how different merge criteria affect the quality of the segmentation and the processing time. The experiment showed that the algorithm automatically computes segmentation thresholds based on local feature analysis. The main limitations of this algorithm were the limited applicability of the adaptive homogeneity tests on very small regions and the order dependency of its segmentation results. The order dependency refers to the fact that different seeds cause different region-growing results.

2.3.3 Clustering

A cluster is a group of similar data. Clustering analysis performs effectively for image segmentation. It allows the partitioning of image data into meaningful regions. Generally, clustering is commonly used in unsupervised segmentation. The difficulties of cluster analysis are related to how many clusters are best and how to determine the validity of the clusters.

Frigui and Krishnapuram [12] described a method called Robust Competitive Agglomeration (RCA) to solve three major issues associated with conventional partition

clustering. The three issues were sensitivity to initialization, difficulty in determining the number of clusters, and sensitivity to noise. RCA starts with a large number of clusters to reduce the sensitivity to initialization and determines the actual number of clusters by a process of competitive agglomeration. The results showed that RCA can provide robust estimates of the number of clusters even when the clusters vary significantly in size and shape, and data are noise contaminated. RCA is not suitable for our project since the histogram of a region can not be estimated accurately if the region is very small. The reason is that histogram estimation is a distribution estimation method. A very small sample size might not result in accurate histogram estimation.

Ng [29] presented an extension to the conventional k-means clustering algorithms by modifying the splitting rule in order to control the number of the clusters. The author developed an approach that allows the use of a k-mean paradigm to efficiently cluster data sets with a fixed number of elements in each cluster by adding suitable constraints.

2.3.4 Morphological watershed

The morphological watershed algorithm is a very powerful automatic segmentation method. It treats a 2-D image as a 3-D model. The $x-y$ spatial coordinates and the intensity of the pixels form the three coordinates of the model. Usually, the input for the watershed is the gradient map of the original image. The morphological watershed algorithm performs well when the gradient map is very clear. However, the watershed

method is very sensitive to noise. This causes an over-segmentation problem which is the main drawback of the watershed algorithm. Some researchers proposed several ways to overcome this problem. Soille [37] introduced the morphological H-minima transform, which modifies the gradient surface, suppressing shallow minima.

Another well-known suppression method is the waterfall method proposed by Beucher [3]. It considers the boundary of a region and then sets the whole region as the minimum value of the region's boundary. Most of the time, local minima suppression is not enough. The image is still oversegmented. Since the input for the watershed is the gradient information of the original image, some pre-processing can be done before the image is passed to the watershed algorithm. Shafarenko *et al.* [36] used a modified gradient map as the input for the watershed algorithm in randomly textured color images. This modified gradient provided both a color similarity measure and a basis for applying the watershed transform. After the watershed transform, a merge process with a termination criterion based on graph theory is used to get the final segmented images. The author concluded that this method is suitable for the automatic processing of granite or any other blob-like image, because it can be fully automatic and it does not require any fine tuning of parameters.

Grau *et al.* [14] presented an improved watershed transform that enables the introduction of prior information in its calculation. Besides the oversegmentation problem, the watershed transform still has other drawbacks: poor detection of significant areas with low contrast boundaries, and poor detection of thin structures. The proposed improvement combines a set of lower-cost functions, one for each of the objects to be

detected in the image. In practical applications, a lot of prior information about the absolute or relative intensities of the objects is available. These lower-cost functions are calculated from the probability values for each voxel and each class. For applications where atlases are available, another improvement was presented [14], using the atlas to generate markers for the watershed transform. The outcome exhibits convincing accuracy for the two applications on which it was tested and the results also suggest the algorithm could be used in many different medical image segmentation problems.

2.3.5 Wavelet coefficients

Wavelets are useful for multiresolution analysis. Moreover, wavelet coefficients are very helpful for image segmentation and texture classification. Many researchers used wavelet coefficients to do feature extraction. Hill *et al.* [16] used a novel marker-location algorithm to locate significant homogeneous textured or non-textured regions. Then, a marker-driven watershed transform was used to properly segment the identified regions. In this method, the term "texture gradient" was introduced. In order to integrate an adaptive scheme with the texture feature extraction process, the authors developed the Non-Decimated Complex Wavelet Packet Transform (NDXWPT). The magnitudes of the coefficients of each complex subband were used to characterize the texture content. Moreover, the authors compared the new marker-location method and traditional methods of marker extraction such as large-scale low-pass filtering [4] or scale-space morphological filtering [17]. These traditional

methods often move or remove salient, small-scale gradient elements that can be vital for effective segmentation. The conclusion was reached that the combined algorithm produced effective texture- and intensity-based segmentation for the application of content-based image retrieval.

O'Callaghan and Bull [30] proposed a two-stage method, which is capable of processing both textured and non-textured objects in a meaningful fashion for general image segmentation. Instead of using NDXWPT to extract the texture's features, the authors used the Dual-Tree Complex Wavelet Transform (DT-CWT) to do texture-feature extraction. Compared to NDXWPT, DT-CWT has its own new features: 1) the computational complexity is greatly reduced since DT-CWT is a decimated, rather than non-decimated, wavelet; 2) orientation-adaptive median filtering is applied to post-processed texture features; 3) a gradient-basin depth criterion is used to control watershed over-segmentation; 4) finally, spectral clustering, based on the novel weighted-mean-cut algorithm, provides a powerful way to incorporate the maximum amount of information about image statistics into the final segmentation. The experimental results indicated that this proposed method is suitable for both textured and non-textured image segmentation.

2.4 Volume calculation

The final goal of the thesis is to measure the changes in ventricular volume. Basically, there are two ways to achieve this goal: 1) calculate the area of the ventricle in each slice, then multiply the area by the thickness of the slice, and, add all the slices which

include the ventricular volume together; 2) construct a 3D model using the whole stack of images, then segment the ventricles in 3D space, and finally, compute the segmented part's volume in 3D space.

Sun [39] used the first method to compute the ventricular volume in her thesis. First of all, segmentation was applied to each image that includes the ventricles in one stack. Then, the volume of each slice was computed. At the last step, all the slices including the ventricles were added together to get the volume. She also proposed a method to reduce the partial-volume effect when calculating the volume. First, choose a region which is exactly inside of the ventricle and then choose another region which exactly includes the whole ventricle. An adaptive threshold was calculated from these two regions and used to decide if a pixel was a ventricle pixel or a boundary pixel. In this thesis, Sun's volume calculation method was used because the segmentation algorithm was based on 2D images.

2.5 Summary

This chapter is a brief literature review of the techniques used in this thesis: wavelet multiresolution analysis, medical image registration, image segmentation, and volume calculation. In this thesis, a framework is proposed to calculate the volume change of the ventricular system in CT images. Some adaptive algorithms are proposed to solve particular problems. A mutual information-based registration algorithm is modified by combining it with multiresolution analysis to correct 3D displacement of the human head between two scans. This adaptive registration has a smoother

mutual information function during optimization than pure mutual information does. A modified watershed segmentation and a novel feature extraction method based on DT-CWT coefficients is introduced to segment the ventricular system with minimal user interaction. This segmentation is not only based on pixel intensity measures, but also on texture measures.

Chapter 3

Theory

Computer vision techniques have been applied to medical-image processing in recent years [13]. In this chapter, the theory behind the computer vision techniques used in this thesis is given. These techniques are: 1) wavelet analysis; 2) image registration; 3) image segmentation; 4) the Kolmogorov-Smirnov test; and 5) volume calculation. These techniques form the fundamental basis of the project. The goal of the project is to calculate the volume change of the ventricular system given two sets of CT images taken at different times.

3.1 Wavelet Analysis

3.1.1 Multiresolution Analysis

The wavelet transform (WT) is a mathematical transformation, much like the Fourier Transform (FT) [13]. They both transform a signal from one domain to another

domain. In digital image processing, the FT is a transformation between space and frequency domains and the WT is the transformation between space and the wavelet domain. Unlike the FT, the WT is more suitable for analyzing "small areas" and provides a good analysis of the original signal at different resolutions. It constitutes a tool to decompose, analyze, and synthesize functions, with an emphasis on space-wavelet localization [13]. For $j, k \in Z$, a simple wavelet decomposition of a function f of space v can be written as the expansion:

$$f(v) = \sum_{j,k \in Z} c_{j,k} \psi_{j,k}(v). \tag{3.1}$$

In Equation (3.1), the functions $\psi_{j,k}$ are the wavelets. They are generated by scaling and translating the "mother" function ψ, in the following way:

$$\psi_{j,k}(v) = 2^{\frac{j}{2}} \psi(2^j v - k). \tag{3.2}$$

Usually, ψ can be expressed by the equation:

$$\psi_{a,b}(v) = \frac{1}{\sqrt{a}} \psi(\frac{v-b}{a}). \tag{3.3}$$

The scale factor a and the translation factor b affect the shape of the children wavelets. According to the definition of wavelets, the continuous wavelet transform [13] is:

$$CWT_x^{\psi}(\tau, s) = \Psi_x^{\psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int f(v) \psi(\frac{v-\tau}{s}) dv. \tag{3.4}$$

As seen in Equation (3.4), the transformed signal is a function of two variables, τ and s, the translation and scale parameters, respectively. $\psi(v)$ is the transforming function, and is called the "mother wavelet". The term "mother" implies that the functions can be used as a basic function to generate children wavelets by scaling and translating it. They are derived from one main function, or the "mother wavelet". The wavelet transform is used to transform the signal $f(v)$ to its wavelet domain. The original signal $f(v)$ can be reconstructed by using the same wavelet. The inverse wavelet transform [13] is defined as:

$$f(v) = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(a,b)\psi_{a,b}(v) \frac{dadb}{a^2}. \tag{3.5}$$

Wavelet analysis is a multiresolution analysis. Multiresolution refers to the representation of a signal at several different resolutions. The most common multiresolution wavelet decomposition is shown in Figure 3.1.

In Figure 3.1a, an image is decomposed into four sub-images. In the four sub-images, subband I_1 is the low-resolution approximation of the original image and the other three subbands D_i represent directional detail information. Figure 3.1b shows the multiresolution analysis applied to a sample image.

An important imaging technique with ties to multiresolution analysis is subband coding. In subband coding, an image is decomposed into a set of band-limited components, called subbands, which can be reassembled to reconstruct the original image without error by using the set of band-limited components. Since the bandwidth of the resulting subbands is smaller than that of the original image, the subbands can

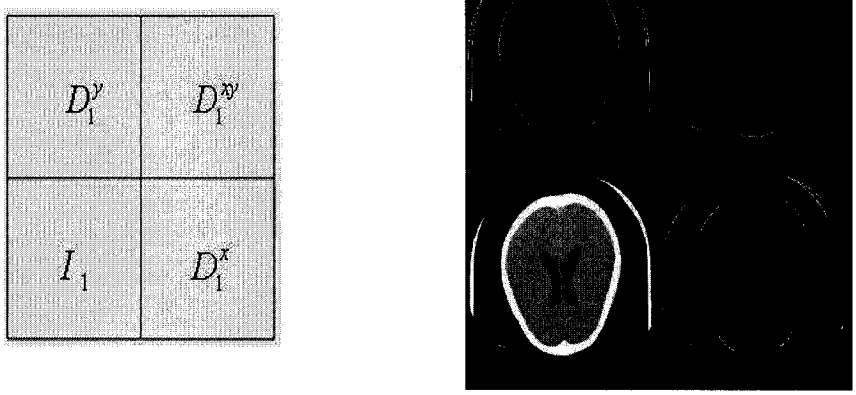

Figure 3.1: Wavelet multiresolution decomposition. From left to right (a-b): a. Multiresolution description: I_1 is the low resolution; D_1^x, D_1^y, D_1^{xy}: Subbands in the horizontal, vertical, diagonal direction, respectively; b. Multiresolution analysis on a medical image

be down-sampled without loss of information. Reconstruction of the original image is accomplished by up-sampling, filtering, and summing the individual subbands. The decomposition algorithm is given by Figure 3.2.

In Figure 3.2, LL refers to the low resolution image and HH, HL, LH are the wavelet subbands which include the detail information along different directions of the image.

Usually, more than one level of decomposition is used to obtain the desired lower resolution. Multilevel decomposition might be needed to decompose the original image more than one level. The multilevel wavelet decomposition is an extension of 1-level decomposition.

The reconstruction of the medical image is exactly the inverse of the decomposition. So, the 2D reconstruction algorithm is illustrated in Figure 3.3.

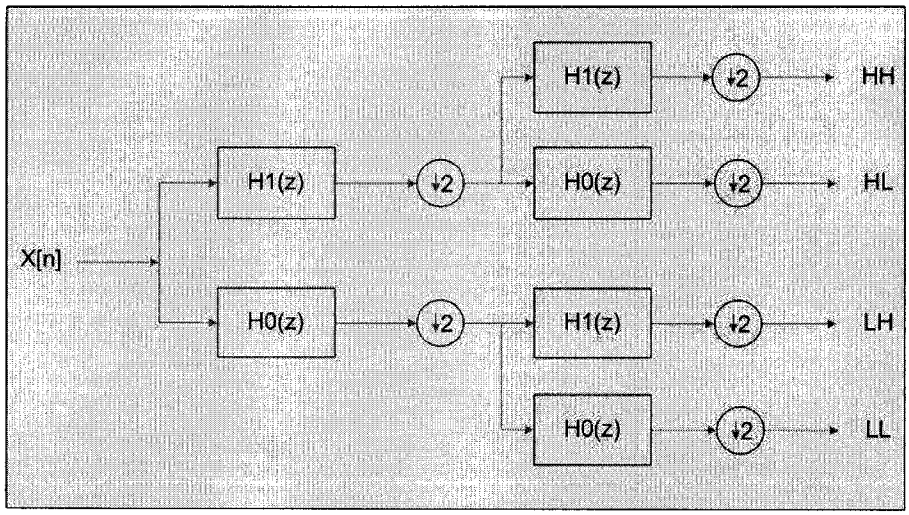

Figure 3.2: 2D Wavelet Decomposition. $X[n]$: the original signal; $H0(z)$, $H1(z)$:lowpass, highpass filter; $\downarrow 2$: downsampling by a factor of 2.

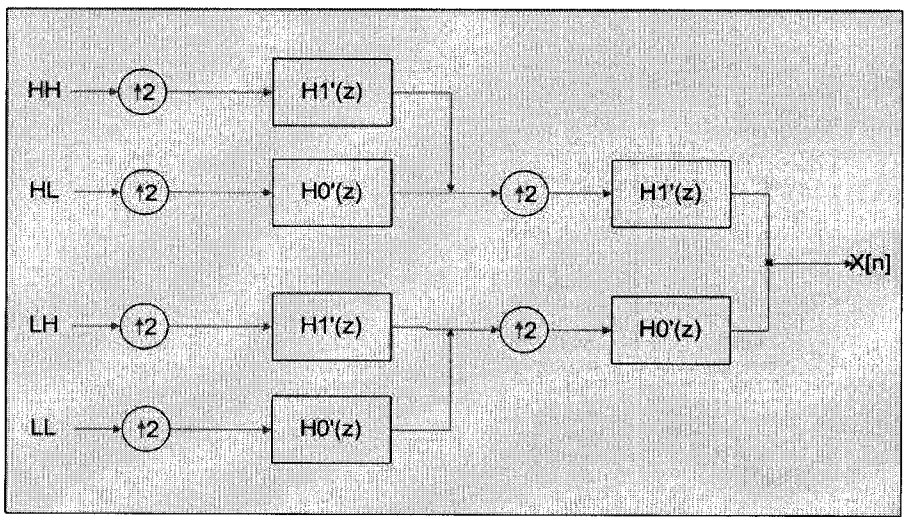

Figure 3.3: 2D Wavelet Reconstruction. LL: lower resolution; HL, LH, HH: wavelet subbands; $\uparrow 2$: upsampling by a factor of 2.

3.1.2 Wavelets in Texture Classification

In addition to the use in multiresolution analysis, wavelet coefficients are helpful for texture classification. Texture classification means to cluster different textures based on their different features. Good texture representation leads to good classification results. Because of the multiresolution characteristic of wavelets, the images are analyzed at more than than a single scale. Moreover, wavelets also provide orientation information for textures. That is why wavelet transforms are found useful to extract new features. The common method for using wavelet coefficients to do feature extraction is energy-based [34]. In this method, the energies of the wavelet coefficients at each scale are calculated for each subband. Then, the mean value and the standard deviation of the energies at each scale and each subband are derived from the image as the components of texture features. Therefore, a texture signature of an image is given by a feature vector which includes the mean values and standard deviations of the wavelet coefficient energies at each scale for each subband. Then a distance measure, such as the normalized Euclidean distance, is chosen to measure similarity between the signatures. This common energy-based method can extract features at more than one scale and serval orientations. However, it suffers from some drawbacks which means modifications were necessary for this project. This texture classification technique assumed the whole image was a homogenous texture and that the image size was a power of two, but in our case, an image has many textures and the textures are in irregular shapes. Therefore, it is very hard to use wavelet coefficients to classify the different textures using the common energy-based method. As discussed

in Chapter 4, a new feature extraction method is proposed to classify the different textures in a single image based on the preliminary segmentation results from the watershed algorithm and the Dual-Tree Complex Wavelet Transform coefficients.

3.1.3 Dual-Tree Complex Wavelet Transform

Although wavelet decomposition provides a good way to represent the features of the signal, it suffers from two main problems which hamper its use for many image analysis and reconstruction tasks. One problem is the lack of shift invariance and the other is the poor directional selectivity for diagonal features [23]. The lack of shift invariance means that small spatial shifts can cause major variations in the distribution of energy between wavelet transform coefficients at different scales. This problem causes energy variations at different scales when using the wavelet coefficients to extract the energy feature from the image at the current resolution [24]. So, the same features might have significant differences due to small shifts of the signal.

The poor directional selectivity for diagonal features is caused because wavelet filters are separable and have real coefficients. In common wavelet decompositions, there is only one subband that represents diagonal image features for each scale. These diagonal features might consist of several sub-diagonal features, such as $\pm 15°$ features, $\pm 45°$ features, and $\pm 75°$ features. If a texture has strong 15 degree features and 75 degree features, common wavelet decomposition will combine them rather than separate these two strong directional features.

Kingsbury [24] proposed a new wavelet-transform, called the Dual-Tree Complex

Wavelet Transform (DT-CWT), to overcome these two main problems (shift variance and poor directional select sensitivity). The DT-CWT is the feature extraction method used in this thesis. DT-CWT is based on the Complex Wavelet Transform. The structure of the Complex Wavelet Transform is the same as that of the Wavelet Transform, except that the filters in the Complex Wavelet Transform have complex coefficients and generate complex output samples. The output sampling rates are unchanged from the wavelet transform. Because of the unchanged output sampling rates and the fact that each sample contains a real and imaginary part, a redundancy of $2:1$ is introduced in one dimension and $4:1$ in two dimensions. If suitable filters for the Complex Wavelet Transform are designed, the magnitudes of their step responses are approximately shift invariant, but the phases vary rapidly. Furthermore, the Complex Wavelet Transform provides a good and true directional selectivity since all the filters and outputs are complex coefficients. For each scale level, more subbands can be generated by the Complex Wavelet Transform than by the Wavelet Transform. For example, a 2-D Complex Wavelet Transform produces six bandpass subbands of complex coefficients at each level, which are strongly oriented at angles of $\pm 15°$, $\pm 45°$, and $\pm 75°$ [24].

3.2 Registration

Registration is the process of finding an optimal transformation to match the information in one image to that in another image. The purpose of registration in this thesis is to minimize the effect of the displacement problem [39]. Displacement of the

human head from one CT scan to the next is an occurrence which is hard to avoid. The displacement problem makes it difficult to compare two sets of CT images which are taken of the same patient but at different times.

Mutual information has been applied to medical image registration because it provides a measure of how similar two images are. In the mutual information registration discussed in this thesis, floating image refers to an image in scan 2 which was taken at time t_2. Reference image refers to an image in scan 1 which was taken at time t_1. Registered image refers to the image obtained by registering the floating image to the reference image.

Pure mutual information usually does not perform well enough for this application because of the substantial noise and the complex movement of the scanned organs in the image. So other techniques, such as interpolation and gradient information, are combined with pure mutual information to improve the performance of mutual-information registration. Interpolation is used when rotating and translating an image. Gradient information can be combined with pure mutual information to get a better measurement [32]. The following subsections introduce the background of mutual information, image interpolation, and gradient information.

3.2.1 Mutual Information

Mutual information is calculated by using Shannon entropy [27]. The definition of Shannon entropy (H) is:

$$H = -\sum_{i=0}^{N-1} p_i \log p_i \tag{3.6}$$

where p_i is the probability of a given symbol. In this respect, the measure is viewed as a measure of uncertainty. The mathematical representation of the mutual information, I, in terms of the entropy, H, between two images A and B is given by [27]:

$$I(A,B) = H(A) + H(B) - H(A,B) \tag{3.7}$$

where $H(A)$ and $H(B)$ are the Shannon entropy for image A and image B, respectively. $H(A,B)$ is the joint entropy between A and B. It denotes the probability distribution of the image intensities shown in both the images A and B. The required entropies $H(A)$, $H(B)$, and $H(A,B)$, can be computed by estimating the probability distribution of the image intensities [27]:

$$\begin{aligned} H(A) &= -\sum_{a} p_a \log p_a \\ H(B) &= -\sum_{b} p_b \log p_b \\ H(A,B) &= -\sum_{a,b} p_{a,b} \log p_{a,b}. \end{aligned} \tag{3.8}$$

In practical calculations, the marginal distributions ($H(A)$ and $H(B)$) are obtained by summing over the rows and the columns, respectively, of the joint entropy histogram, so only the joint histogram needs to be constructed.

Mutual information has some useful properties [27] such as:

(i) $I(A,B) = I(B,A)$. Mutual information is symmetric; otherwise it would not be mutual information. However, this symmetry is an approximation since implementational aspects of a registration method, such as interpolation and number of

samples, can result in differences in outcome when registering A to B or B to A.

(ii) $I(A, A) = H(A)$. The mutual information between image A and itself equals Shannon Entropy of image A.

(iii) $I(A, B) \leq H(A)$ and $I(A, B) \leq H(B)$. The mutual information between two images A and B can not be greater than the information of image A and image B themselves.

(iv) $I(A, B) > 0$. The uncertainly about A cannot be increased by learning about B.

Now, the registration criterion is easily obtained by using the properties of mutual information. Assume A and B are two images which are related by the registration transformation T with parameter vector α, such that pixel P in A with intensity a corresponds to pixel $T(P)$ in B with intensity b. Using properties (ii) and (iii) above, it can be found that when the mutual information value is maximum, image A and image B are the most similar [27]. Therefore, the mutual information registration criterion requires that for the images to be geometrically aligned by the transformation $T(\alpha)$ for which $I(A, B)$ is maximal:

$$\alpha^* = arg \max I(A, B) \tag{3.9}$$

where α^* is a parameter vector which is the optimal parameter vector obtained in equation (3.9). However, the optimal parameter vector α^* can differ from the perfect parameter vector. Assuming the perfect transformation parameter vector is α_p, then, the misregistration is $|\alpha_p - \alpha^*|$. In order to reduce the misregistration rate, $I(A, B)$

should vary smoothly as a function of the misregistration. This requires p_a, p_b, and $p_{a,b}$ to change smoothly when α is varied [27].

A rough mutual information function will cause overlap between two images. Overlap refers to the problem that with increasing misregistration, the mutual information measure may actually increase. This can occur when the relative areas of object and background even out and the sum of the marginal entropies increases faster than the joint entropy. Studholme [38] found that normalized mutual information is less sensitive to change in the overlap between the images. Therefore, $NMI(A, B)$ is used in this thesis:

$$NMI(A, B) = \frac{H(A) + H(B)}{H(A, B)}. \tag{3.10}$$

3.2.2 Image Interpolation

Image interpolation is the process of determining the values of the image function at positions that lie between its samples. It assumes that there is a continuous function through the discrete input samples. The considered values can be in arbitrary positions between the input samples. The image quality highly depends on the interpolation method used for processing. The interpolation techniques are divided into two categories, deterministic and statistical interpolation techniques [13]. Since statistical interpolation is usually inefficient, deterministic interpolation was applied in this thesis. Deterministic interpolations assume certain variability between the sample points, such as linearity in the case of linear interpolation. Two popular de-

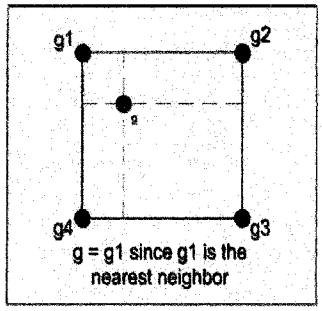

Figure 3.4: Nearest neighbor interpolation

terministic interpolation methods are discussed and the methods are compared with each other based on the interpolation results.

Nearest Neighbor Interpolation

Nearest neighbor interpolation is the simplest method of interpolation. It finds the nearest sample point to the desired point and then assigns the desired point with the nearest sample point value. It is also called replicated interpolation. The two-dimensional nearest neighbor interpolation is an extension of the one-dimensional nearest neighbor interpolation. Figure 3.4 gives a good demonstration of the 2D nearest neighbor interpolation.

Linear Interpolation

Linear interpolation is another common deterministic interpolation method. It assumes a linear relationship between all the input samples. Linear interpolation calculates the desired values based on a linear relationship rather than simply repli-

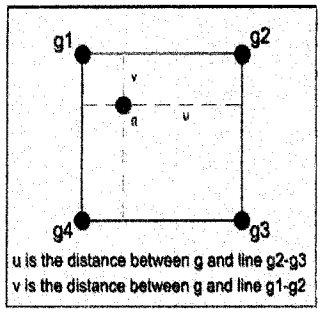

Figure 3.5: Bilinear interpolation

cating the nearest input samples as nearest neighbour interpolation does. The one-dimensional linear interpolation of pixel grey level, $g(x)$ at position x is given by:

$$g(x) = \frac{x_1 - x}{x_1 - x_0} g_0 + \frac{x - x_0}{x_1 - x_0} g_1 \tag{3.11}$$

where g_0 and g_1 are two pixel gray levels at positions x_0, and x_1, respectively. Bilinear interpolation is two 1D linear interpolations along the x axis and the y axis, respectively. It considers four neighbor pixels and can interpolate any pixel value inside of the rectangle which is formed by these four neighboring pixels. This is shown in Figure 3.5.

Using the distances u and v as defined in Figure 3.5, the function $f_{bilinear}()$ is

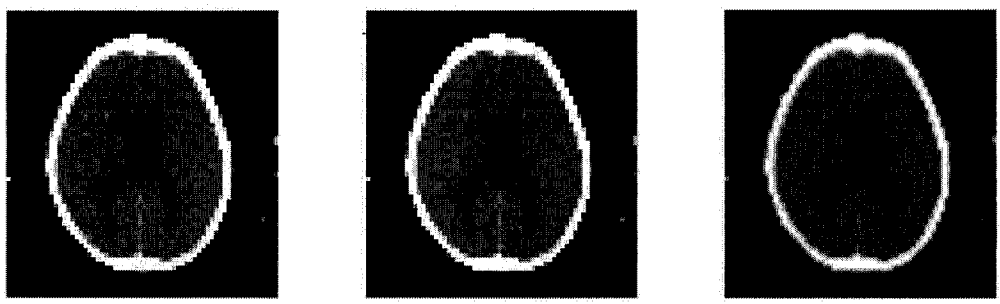

Figure 3.6: Interpolation comparison. From left to right (a-c): a. Original image (64 $\times$ 64); b. Nearest neighbor interpolation (128 $\times$ 128); c. Bilinear interpolation (128 $\times$ 128)

defined as [13]:

$$g = f_{bilinear}(g_1, g_2, g_3, g_4)$$
$$= g_1 * u * (1-v) + g_2 * (1-u) * (1-v)$$
$$+g_3 * (1-u) * v + g_4 * u * v. \tag{3.12}$$

These two deterministic interpolations have their own advantages and disadvantages. Nearest neighbor interpolation is the simplest of the two methods but it produces jagged edges in the image. Linear interpolation produces smoother results than nearest neighbor, however, it blurs the image. Figure 3.6 compares the two interpolations' results by resampling an image to double its size. Linear interpolation was used in both the registration and segmentation algorithms. In registration, linear interpolation was applied to the inside of the slices and between slices.

3.2.3 Gradient Information

The gradient is commonly used as the measure of slope. It represents the inclination of a surface along a given direction. In mathematics, gradient refers to the first derivatives of a function. The general mathematical formula for the gradient in 2D is [13]:

$$\nabla F = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \end{bmatrix} \tag{3.13}$$

where G_x and G_y are the gradients along the x and y directions, respectively. Then, the magnitude ∇f and the phase $\nabla \rho$ of ∇F are given by:

$$\begin{aligned} \nabla f &= |\nabla F| \\ &= \sqrt{[G_x^2 + G_y^2]} \\ &= \sqrt{\left[\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2\right]} \end{aligned}$$

$$\nabla \rho = \tan^{-1}\left(\frac{\partial F}{\partial y} / \frac{\partial F}{\partial x}\right). \tag{3.14}$$

Using Equations (3.14), the gradients in an image can be computed. First, the component of the gradient along the x axis is calculated followed by the component along the y axis. Finally, the magnitude and the phase, which form the gradient information of the image, are obtained. Gradient information is helpful for some further image processing, such as edge detection and the morphological watershed algorithm. In digital image processing, the gradient map is usually obtained by applying some gradient operators to the images in the spatial domain. Two main gradient operators, Sobel and Gaussian, are introduced. Both these two operators are popular in image

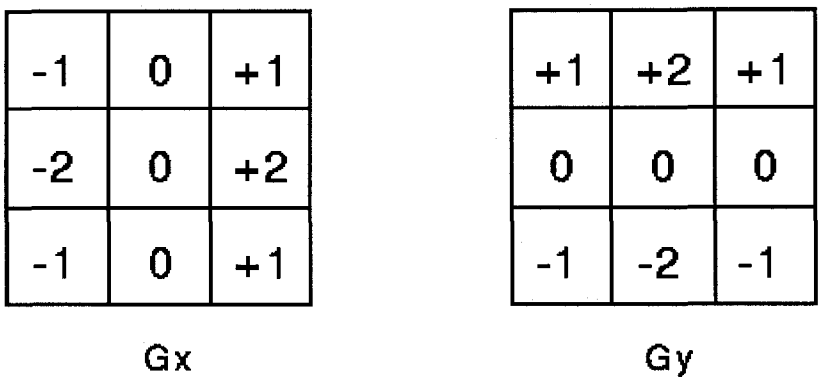

Figure 3.7: Sobel mask. From left to right (a-c): a. Gradient along x axis; b. Gradient along y axis

processing and were used in this thesis.

Sobel Gradient Operator

The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial gradient that correspond to edges [13]. Typically it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. A pair of convolution masks of the Sobel operator is shown in Figure 3.7. The pair of masks were used in the registration algorithm to accurately detect the gradient information of each pixel. The smallest mask size 3×3 was chosen.

The masks are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid, one mask for each of the two perpendicular orientations. In the processing, the masks are passed over the input image separately, to produce separate measurements of the gradient components in each orientation,

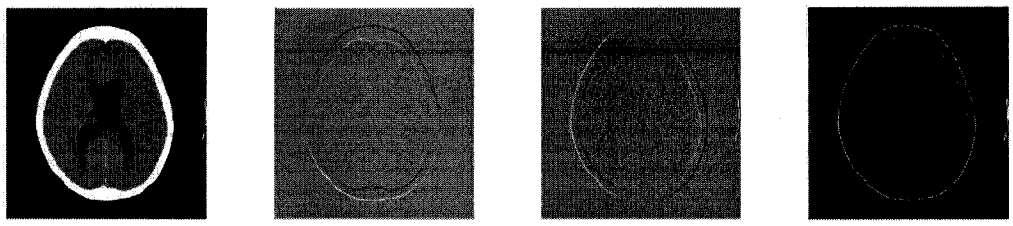

Figure 3.8: Sobel gradient. From left to right (a-d): a. Original image; b. Gradient along x axis; c. Gradient along y axis; d. Gradient magnitude of the image

G_x and G_y (see Figure 3.7). The gradient components can then be combined to find the magnitude (Equation(3.2.3)) and orientation (Equation(3.14)) of the gradient at each point. The result of applying the Sobel operator to an example image is shown in Figure 3.8.

In this thesis, the Sobel gradient operator was used in the registration algorithm since pixel gray levels were the values of interests when computing the mutual information between two images.

Gaussian Gradient Operator

The Gaussian gradient operator is another operator and is based on the derivative of the Gaussian function. It is designed to smooth an image and suppress the noise before detecting the gradients [13]. The operator effectively convolves the image with a Gaussian kernel of width σ, the standard deviation, and then takes the first order derivative of the function. The operator, $\bigtriangledown F$ can be defined:

$$\bigtriangledown F = g'(x,y) = G'_\sigma(x,y) * f(x,y) \tag{3.15}$$

where

$$g(x,y) = G_\sigma(x,y) * f(x,y) \tag{3.16}$$

and

$$G_\sigma(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{x^2+y^2}{2\sigma^2}\right).$$

$G'_\sigma(x,y)$ is the first order derivative of $G_\sigma(x,y)$. The Gaussian kernel of any size is obtained by approximating the continuous expression of the Gaussian given above. However, the sum (or average) of all elements of the kernel has to be zero so that the convolution result of a homogeneous region is always zero. Compared to the Sobel gradient operator, the Gaussian gradient operator smoothes the image first and then finds the gradient information based on the smoothed image. In this thesis, the Gaussian gradient operator was applied in the segmentation algorithm with the standard deviation $\sigma = 2.0$. It was appropriate for that application because it was necessary to suppress the gradient information inside a region with homogeneous texture and only detect the edges of the texture regions. Figure 3.9 shows an example of applying the Gaussian operator to an image. The Sobel operator is applied to the same image for comparison.

Figure 3.9: Comparison between Gaussian gradient and Sobel gradient results. From left to right (a-c): a. Original image; b. Gaussian gradient with $\sigma = 5$; c. Sobel gradient with mask size 3×3

3.3 Segmentation

3.3.1 Denoising

Usually, the raw image data is unsuitable for segmentation because the data includes noise that might significantly affect the segmentation results. Some pre-processing steps have to be applied before segmentation. Denoising is one of the most important steps during the segmentation process. Two common denoising filters were used in this thesis: the Gaussian filter and the median filter.

Gaussian Filtering

The Gaussian lowpass filter (GLPF) is one of the most popular smoothing filters. The GLPF can be defined in both the spatial and frequency domains. The formula of the GLPF in the frequency domain [13] is:

$$H(\mu, \nu) = e^{\frac{D^2(\mu,\nu)}{2*\sigma^2}} \tag{3.17}$$

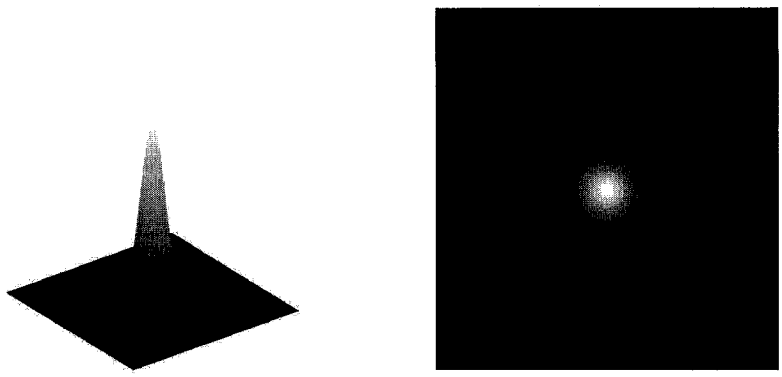

Figure 3.10: Gaussian lowpass filter. From left to right (a-b): a. 3D visualization of Gaussian Lowpass Filter; b. Gaussian Lowpass Filter displayed as an image

where $D(\mu, \nu)$ is the distance from the origin of the Fourier transform and σ is a measure of the spread of the Gaussian curve. μ and ν are the 2D coordinates in the frequency domain. The shape of the Gaussian is determined by the value of σ, and σ is also referred to as the cutoff frequency. Different σ will cause different distributions of Gaussian lowpass filter and hence different smoothing results. Figure 3.10 is a visualization of the Gaussian lowpass filter in the frequency domain.

The Gaussian lowpass filter has a good denoising effect, but as the cutoff frequency σ is decreased, the image becomes blurred. A tradeoff can make the signal-to-noise ratio (SNR) as large as possible. Figure 3.11 shows the denoising results with different cutoff frequencies.

Median filter

The median filter is another prevalent smoothing filter which not only smoothes the image data but also has good edge preservation. Its performance is particularly good

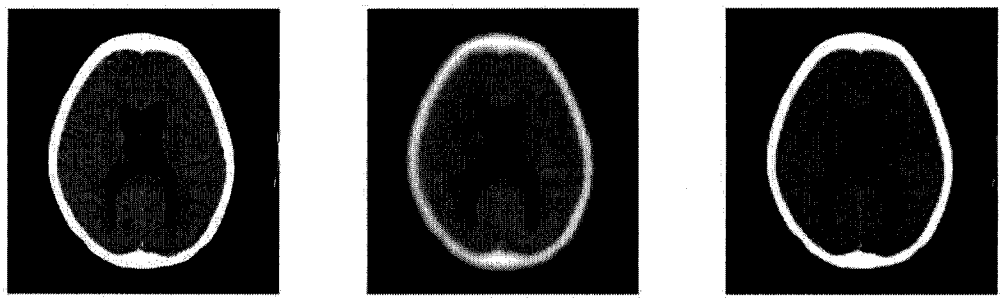

Figure 3.11: GLPF denoising. From left to right (a-c): a. Original Image; b. Denoising with $\sigma = 15$ in frequency domain; c. Denoising with $\sigma = 60$ in frequency domain

for removing shot noise [13]. However, in this project, the median filter is used to get rid of the detail information inside a texture region while keeping the edges between different textures regions. Median filtering is applied in the spatial domain. Once a suitable mask size is determined, the mask is passed over the image and all the pixels which are covered by the mask are selected. Then all the selected pixels' gray values are sorted into ascending order. Now, the median value of this list is output as the pixel gray value. The effect of median filtering is shown in Figure 3.12. The median filter performs better at preserving the edges in the images than the GLPF (Figure 3.11) does.

3.3.2 Mathematical Morphology

Common morphological operations

Mathematical morphology is the study of form and structure. In digital image processing, morphology uses set theory from mathematics to process images. Generally,

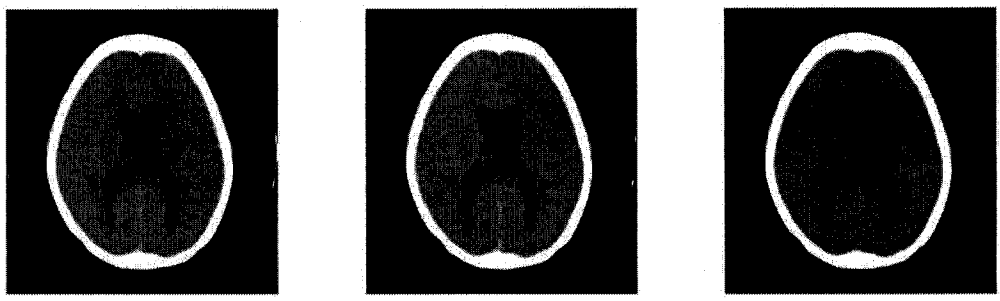

Figure 3.12: Median Denoising. From left to right (a-c): a. Original Image; b. Denoising with 5×5 mask; c. Denoising with 15×15 mask

morphological operations are performed on binary images (images with only two grey levels). However, many morphological operations are modified to apply to grey-level images. There are many morphological operations, such as dilation, erosion, opening, closing, and watershed. The theory of erosion and watershed will be given more weight since these two morphological operations were applied in this thesis.

First of all, a short summary of set theory will be introduced since the theory is the fundamental basis for morphological operations [13].

1. If you have two sets, S_1 and S_2, their union is the set of all the elements that belong to S_1 and all the elements that belong to S_2. When processing two binary images, this can be implemented by merging the images with a logical OR.

2. The intersection of two sets is the set of elements that belong to both S_1 and S_2, but none of the elements belong to S_1 only or S_2 only. When processing two binary images, this can be implemented by merging the images with a logical AND.

3. The difference of S_1 and S_2 $(S_1 - S_2)$ is the set of elements in S_1 that are not in S_2.

4. The complement of a set S_1 is the set of elements that are not in S_1. When processing binary images, this can be implemented by using a logical NOT.

5. Two sets, S_1 and S_2, are disjoint if they have no common elements.

Morphology simply applies these rules to images. Because two sets S_1 and S_2 are usually compared, morphological operations have two common components: a structuring element (SE) and a window element (WE). The structuring element, which is a matrix consisting of only 0's and 1's, and can have any arbitrary shape and size, is used to probe the input images. The pixels with values of 1 define the adjacent neighborhood. The window element includes the image pixel and its adjacent neighbors as defined by the structuring element when the structuring element is passed over the image as a mask.

Dilation is defined as the set of all points where the union of the structuring element and the image is non-empty [13]. In the implementation, the structuring element is passed over the image. For each source pixel, if any of the pixels in the SE are "1" and line up with a source pixel which is also "1", the output pixel is "1". The effect of dilation is to increase an object's geometric area.

Erosion is defined as the compliment of the dilation of the compliment of the image with SE [13]. The mathematical description of the relationship between erosion and dilation is:

$$M_{Erosion}(S_1, SE) = \overline{(M_{Dilation}(\overline{S}_1, SE))} \tag{3.18}$$

Compared to dilation, erosion has the effect of decreasing the object's geometric area.

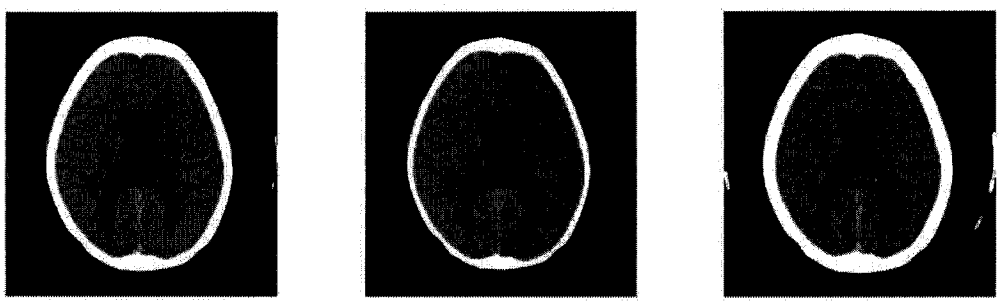

Figure 3.13: Erosion and Dilation. From left to right (a-c): a. Original Image; b. Erosion effect; c. Dilation effect

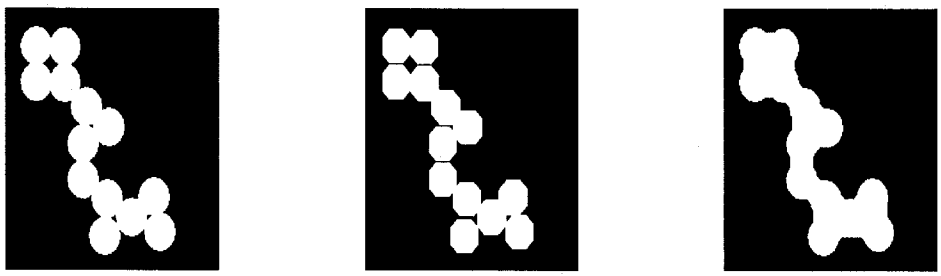

Figure 3.14: Opening and Closing. From left to right (a-c): a. Original binary Image; b. Opening effect; c. Closing effect

The following two images in Figure 3.13 show an image after erosion and dilation.

Dilation and erosion form the basis of morphological image processing. The opening and closing operations are defined based on dilation and erosion. The opening operation is the dilation of the erosion of the image. It tends to smooth outward bumps, break narrow sections and eliminate thin protrusions. On the other hand, the closing operation is the erosion of the dilation of the image. It tends to eliminate small holes and remove inward bumps. Figure 3.14 shows the effect of the opening and closing operations on a binary image.

Morphological watershed transform

The morphological watershed transform was proposed by Digabel and Lantuejoul [13] and later improved by Beucher and Lantuejoul [3]. It can be classified as a region-based segmentation approach. The intuitive idea underlying this method comes from geography: it is that of a landscape or topographic relief which is flooded by water. Watersheds are the divide lines of the domains of attraction of rain falling over the region [13]. The concept of watersheds is based on visualizing an image in 3-dimensions: two spatial coordinates and the gray levels. So, an image is visualized as a mountain. Now, the water is starting to flood the whole mountain from the catchment basins. In this "mountain", three types of points are considered: a) points belonging to a regional minimum; b) points that have already been flooded; c) points at which water would be equally likely to fall to more than one such minimum [13]. The points satisfying condition (c) are termed *divide lines* or *watershed lines.* The principle objective of the watershed segmentation algorithm is to find the watershed lines in an image. The general idea is simple: water floods from the lowest catchment basins, which refer to the lowest gray levels in an image. Each catchment basin is given an individual label. Then, water is allowed to flood from catchment basins at a flooding rate. Now, there are two situations that could happen: (1) two catchment basins could have the same label or two catchment basins could have different labels. If two catchment basins which have the same label are going to merge, these two regions will become one region and keep the same label. If two catchment basins which have different labels are going to merge, a dam has to be built to prevent the

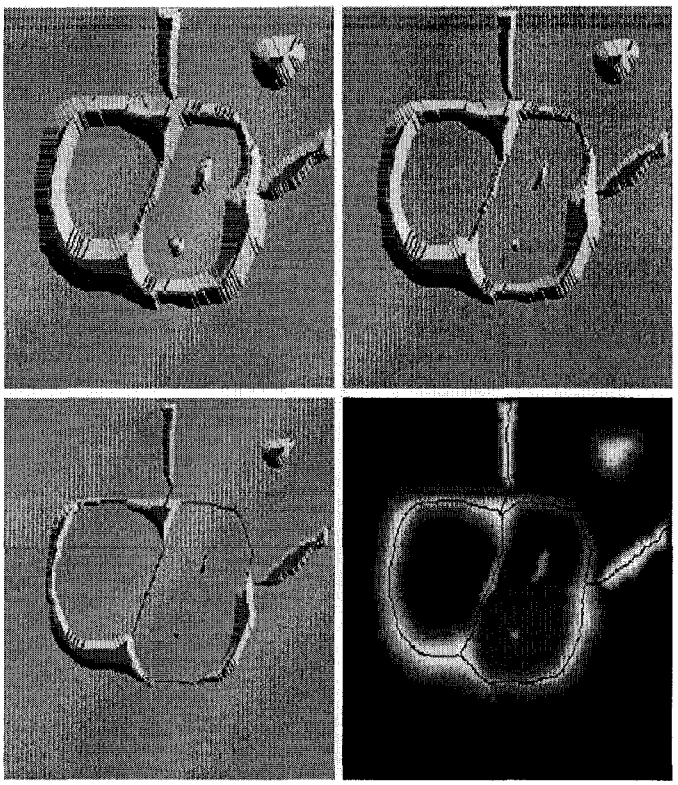

Figure 3.15: Watershed process [13]. From left to right (a-d): a. Topographic view; b. First stage of flooding; c. Middle stage of flooding; d. Final watershed (segmentation)

merging of these two regions. In practical image processing, the gray level of the dam is usually the maximum gray value plus 1. The flooding will eventually reach a stage when only the tops of the dams are visible above the water line. The boundaries of these dams correspond to the boundaries in the image extracted by a watershed segmentation algorithm. The flooding process shows clearly in Figure 3.15.

The algorithm definition of watershed based on immersion was given by Vincent and Soille [42]. The algorithm is implemented by using the following steps.

1. Sort the gray levels in the image in ascending order and start to flood from the lowest gray level.

2. Mark the flooded pixels as the catchment basin. Check if two regions are going to merge together. If the labels of these two regions are different, mark the pixels between these two regions as watershed (or dam) pixels. Otherwise, merge these two regions in the next step.

3. Flood pixels with the higher gray levels. Continue to do step 2.

4. The process stops when the flooding reaches the highest gray level in the image. Now, all the pixels in the image are marked with region labels or watershed labels.

Direct application of the watershed algorithm causes an over-segmentation problem due to noise and other local irregularities of the gradient [13]. As shown in Figure 3.16, over-segmentation has a significant effect on the outcome of the watershed algorithm. In this case, a lot of segmented regions exist in the image. There are several ways to deal with this problem, such as marker-based watershed and H-minima suppression. Marker-based watershed requires the user to specify some markers in the image. This increases the user interaction. So, H-minima suppression was applied in this project. H-minima suppression modifies the gradient surface, suppressing shallow minima by setting a threshold. For example, when the gradient map of an image is obtained, a threshold is assigned to determine all the gray levels below this threshold as the catchment basins. This, very much like the watershed transform, can be imagined as an immersion process: minima are filled with a specified depth of water.

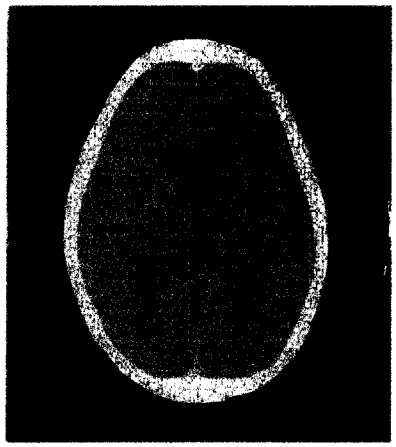

Figure 3.16: Over-segmentation problem in watershed algorithm

3.4 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test (KS-test) is used to determine if two datasets differ significantly. The KS-test has the advantage of making no assumption about the distribution of the data [43] and compares two samples without any parameters. For each potential value x, the KS test compares the proportion of $X1$ values less than x with proportion of $X2$ values less than x, where $X1$ and $X2$ are two observered samples. The KS-test was used in this thesis to compare the distance between two adjacent regions in an image. If two regions have similar probability distributions, the KS test will return a value close to 1, otherwise, it will return a value close to 0. An existing function in Matlab was used to implement the KS-test. The function's name in Matlab Statistical Toolbox is $kstest2(sam_1, sam_2)$, where sam_1 and sam_2 are two samples with finite size. The return value tells whether there is a significant difference between sam_1 and sam_2.

3.5 Volume calculation

After segmentation of the ventricular system, a volume calculation algorithm was applied to get the volume of the ventricular system. There are two very important variables for computation: pixel spacing P_s and slice thickness $S_{thickness}$. First of all, the area of the ventricle in one image is obtained. Then, the area is multiplied by the slice thickness to get the volume of the ventricles in a single image. Finally, all the volumes from the slices which contain the ventricles are added together. If A_i is the area of the ventricles in the current image and V_i is the volume of the ith slice, then:

$$A_i = P_s * P_s * N_{pixel} \tag{3.19}$$

$$V_i = A_i * S_{thickness} \tag{3.20}$$

where N_{pixel} is the number of white pixels in the ith slice. The total volume V is:

$$V = \sum_{i=1}^{N_{slice}} V_i \tag{3.21}$$

where N_{slice} is the total number of slices which contain the ventricles.

Chapter 4

Framework Design and Implementation

In Chapter 3, the fundamental theories used in this thesis were reviewed: wavelet theory, mutual information registration methods, morphological operations, the morphological watershed algorithm, the KS-Test, and the volume calculation algorithm. In this Chapter, a framework based on these theories is presented. This framework provides a quantitive way to measure the volume change of the ventricular system using CT images. Three aspects are discussed:adaptive registration, adaptive segmentation, and volume calculation. The "adaptive registration" section presents a modified method for image registration that combines wavelet multiresolution analysis and mutual information. A novel feature extraction method based on the DT-CWT coefficients is introduced in the "adaptive segmentation" section and the last section describes the method used for volume calculation. Fig 4.1 shows the whole progress

of the framework.

4.1 Adaptive registration

4.1.1 Incorporating Gradient Information

As mentioned in Chapter 3, the goal of mutual information registration is to find the global maximum when optimizing the parameters which form the affine transformation. A mutual information function could be obtained using a variety of different parameter sets. The smoother the function is, the easier it is to find the global maximum. Unfortunately, the mutual information function is very rough. Usually using just mutual information is not enough because a random reshuffling of the image voxels (identical for both images) yields the same mutual information value as for the original image. This means that there are a lot of local maxima which cause misregistration, and then, displacement correction fails. According to Pluim *et al.* [32], gradient information can be combined with mutual information to reduce the local maxima during the mutual information calculation. The method is to multiply the mutual information with a gradient term. The gradient term is based on both the magnitude and the orientation of the gradient.

Because linear interpolation is used between slices, a 2D gradient vector can be used to represent the gradient information for each pixel. Then, the gradient vectors are computed in the floating image and the corresponding points are found in the registered image, by geometric transformation of the reference image. The two partial

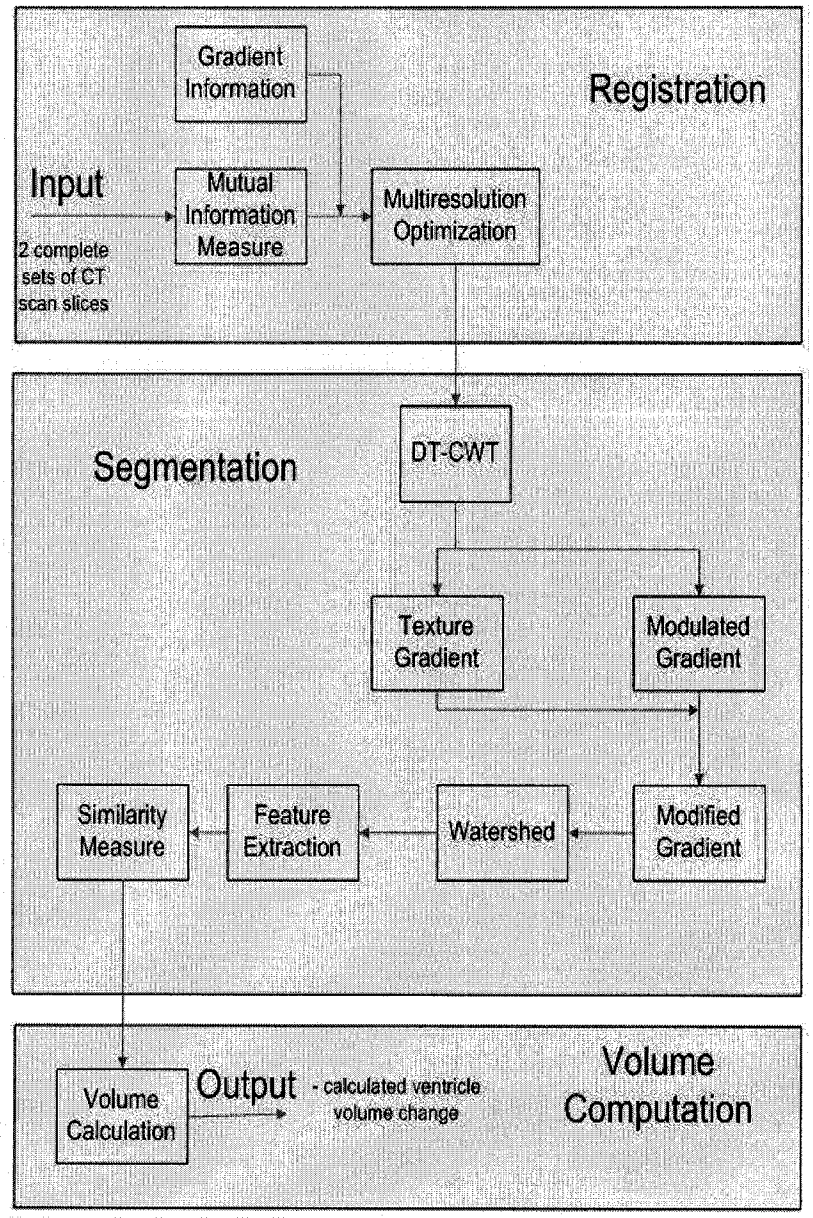

Figure 4.1: Framework

derivatives that together form the gradient vector are calculated by convolving the image with the appropriate first derivatives of a Gaussian kernel of scale σ. So, the angle $\alpha(\sigma)$ between the gradient vectors is defined by:

$$\alpha(\sigma) = \arccos \frac{\bigtriangledown x(\sigma) \bullet \bigtriangledown x'(\sigma)}{|\bigtriangledown x(\sigma)||\bigtriangledown x'(\sigma)|} \tag{4.1}$$

where $\bigtriangledown x(\sigma)$ and $\bigtriangledown x'(\sigma)$ denote the gradient vectors in the floating images and registered images, respectively. $||$ is the magnitude of a vector and $\bullet$ is the convolution. After getting the angle $\alpha(\sigma)$ of two corresponding pixels, a weighting function f_{weight} is used to compute the weight of each pair of corresponding pixels.

$$f_{weight}(\alpha) = \frac{\cos(2\alpha) + 1}{2}. \tag{4.2}$$

According to the Equation (4.2), if the angle is 0° or 180°, the weighting function will have the largest value. In other words, this pixel is a significant pixel which should have a significant effect on the mutual information calculation. Because only strong gradients that appear in both images are considered, the angle function is multiplied by the minimum of the gradient magnitudes. The gradient term is then computed by summing all the resulting products for all pixels. Therefore, the mutual information becomes:

$$I_{new}(A, B) = G(A, B)I(A, B) \tag{4.3}$$

with

$$G(A,B) = \sum_{(x,x' \in (A \cap B))} f_{weight}(\alpha_{x,x'(\sigma)}) \min(|\bigtriangledown x(\sigma)|, |\bigtriangledown x'(\sigma)|). \tag{4.4}$$

Similarly, the new normalized mutual information (NMI) is defined as

$$NMI_{new}(A,B) = G(A,B)NMI(A,B). \tag{4.5}$$

4.1.2 Voxel-based Interpolation

The displacement of the human head is not only in-plane, but also out-of-plane. When the same patient has a CT scan taken at two different times, it is possible that the patient might turn a little at the second scanning. This is in-plane movement. It is also likely that the patient might have a slight nodding action or tilt the head to the side. These are out-of-plane movements because they are across the slices. Since there are both 2D (in-plane) and 3D (out-of-plane) movements, the whole stack of images has to be used for registration. Then, a 3D model can be constructed. Now, a pixel in 2D space becomes a voxel in 3D space because it has volume in 3D space. Since the slice thickness is greater than the pixel spacing, interpolation has to be used between slices as well as between pixels inside of the slice. Voxel-based interpolation is used to map 2D pixel images to 3D voxel images. The size and shape of one voxel is very important for the desired results. A cube with pixel spacing as its length is used as one voxel. The choice of the cube avoids deformation problems during the affine transformation. Trilinear interpolation is used both inter-pixel and inter-slice.

Trilinear interpolation is linear interpolation along the x axis, the y axis, and

the z axis, respectively. Moreover, partial volume interpolation is used to accurately compute the joint histogram using the method described by Maes [25]. Instead of interpolating new intensity values in the registered image, the contribution of the image intensity to the joint histogram, h, is used as a weight, ω_i. The interpolation value which corresponds to x in the reference image is not used, but rather the four neighboring pixels y_i in the reference image, are used. The mathematical representation is:

$$h(F(x), R(y_i)) = h(F(x), R(y_i)) + \omega_i \tag{4.6}$$

where $F(x)$ and $R(y_i)$ are the gray values in the floating image and the reference image, respectively. The ω_i refers to the weight of these four neighbor pixels. Without interpolating new intensity values, partial volume interpolation reduces the local maxima significantly.

4.1.3 Optimization

In order to find the global maximum of the mutual information function, an optimization method has to be used. There are two important properties of the mutual information function that might influence the choice of optimization method. One is the local maxima. We have already mentioned some ways to reduce the local maxima. The other is the range of the optimum. There are six parameters of the affine transformation (3 for 3D rotational angles and 3 for 3D translations) to be optimized simultaneously. These six parameters are not independent since the displacement might be caused simultaneously by translation and rotation. Thus the inverse down-

hill simplex method is used in this thesis. Although it is time-consuming, the simplex method does not require derivatives and can optimize n parameters simultaneously [28].

A simplex is defined as a figure of $N+1$ vertices in the N-dimensional search space. In this case, the vertices are made up from the six transformation parameters (3 for rotation and 3 for translation) and all seven vertices have random initial values. On each iteration, these vertices are used to compute the value of the mutual information. Then, the simplex method compares these values of mutual information and expands, contracts, or reflects the vertices. When the distance between each pair of vertices is smaller than a tolerance, the search process stops. The parameters used for the final iteration are selected as the calculated 3D displacement of the human head.

4.1.4 Incorporating Wavelet Multiresolution

Even though some work had been done to reduce the number of local maxima, it was still a problem during the displacement correction. So, wavelet multiresolution decomposition was also incorporated with the mutual information.

The optimization is performed initially at a lower resolution so only a fraction of the voxels in the image are used to construct the joint histograms. At the lower resolution, detail information is removed, the mutual information function is smoother, and the local maxima are significantly suppressed. Thus it is easy for the simplex method to find the global maximum. After the global maximum is found at the current resolution, the resolution level is increased and initialization is based on the

previously found maximum. Therefore, a combination of mutual information and multiresolution analysis improves the chance of finding the global maxima in the mutual information function.

4.2 Adaptive segmentation

The segmentation algorithm is one of the most important parts in the algorithm framework. The ultimate goal is to measure the volume of the ventricular system and then to calculate if there is any significant change in volume in the time between the two CT scans. An accurate segmentation algorithm contributes to an accurate volume calculation.

In this thesis, an adaptive segmentation based on texture measures and the watershed algorithm is proposed. User interaction is kept to a minimum. In the traditional segmentation framework, a single "best" segmentation is obtained for the classification. The "best" segmentation results refer to the acceptable segmented outcome based on the human visual model. It is usually very hard to get the "best" results. The classification method refers to the clustering algorithms which are used to merge homogenous or non-homogenous regions. However, the adaptive segmentation used in this thesis is not built on the traditional segmentation framework, but rather on a wrapper-based segmentation framework [10]. Compared to the traditional segmentation framework, the wrapper-based framework provides a closed-loop to evaluate the results of segmentation. This closed-loop will not only adapt the parameters of the feature extraction algorithm, but will also actually direct the final segmentation

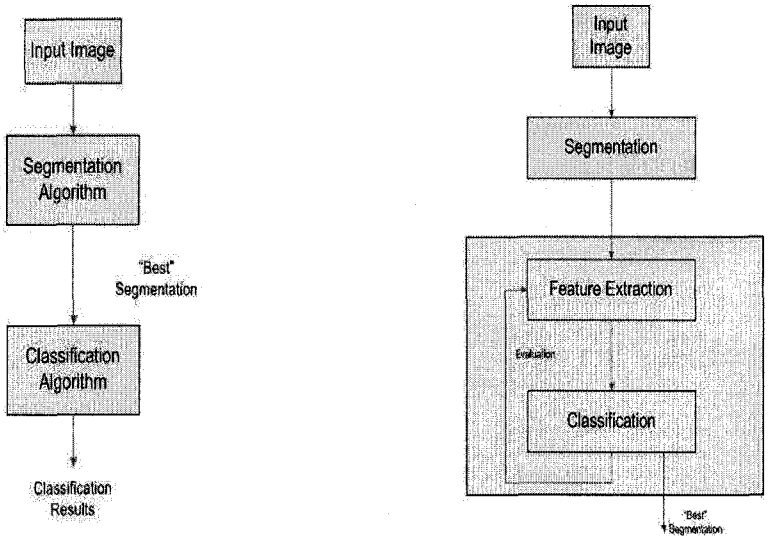

Figure 4.2: Comparison. From left to right (a-b): a. Traditional segmentation; b. Wrapper-based segmentation

based on the underlying shape characteristics of the object of interest. The difference between the traditional framework and wrapper-based framework is shown in Figure 4.2.

4.2.1 Texture Gradient

The whole structure of this adaptive algorithm is based on watershed segmentation. The theory of watershed was already introduced in Chapter 3. As mentioned in the watershed theory, the input to the watershed is the gradient information of the original image. Usually, this gradient information is calculated based on the intensities of each pixel. This simple watershed segmentation causes serious over-segmentation problems [13].

Texture boundaries have been used for the effective partitioning of images [16]. Texture is important because, in theory, different textures contain information that can be used to identify different tissues. Therefore, if the gradient between textures can be detected and used as the input of the watershed algorithm, images can be segmented into several homogenous texture regions. Hill [16] introduced this gradient as the *texture gradient.*

Dual-tree complex wavelet transform representation

In this thesis, the texture gradient is derived from the Dual-Tree Complex Wavelet Transform (DT-CWT) coefficients [23] as dicussed in Chapter 3.1.3. DT-CWT calculates the complex wavelet transform of a signal using two separate discrete wavelet transform decompositions (tree a and tree b). It retains the useful properties of scale and orientation sensitivity and is approximately shift invariant, while also providing a representation with reduced redundancy. For each scale level, six subbands which are oriented at $\pm 15°$, $\pm 45°$, and $\pm 75°$ are produced. In other words, the detail information of the original image is represented along six different orientations. The block diagram for the DT-CWT is shown in Figure 4.3. In this thesis, $D_{i,\theta}(x, y)$ is used to represent the subband image with orientation angle θ at scale level i and spatial co-ordinates x-y. An example of subimages after DT-CWT has been applied is given in Figure 4.4.

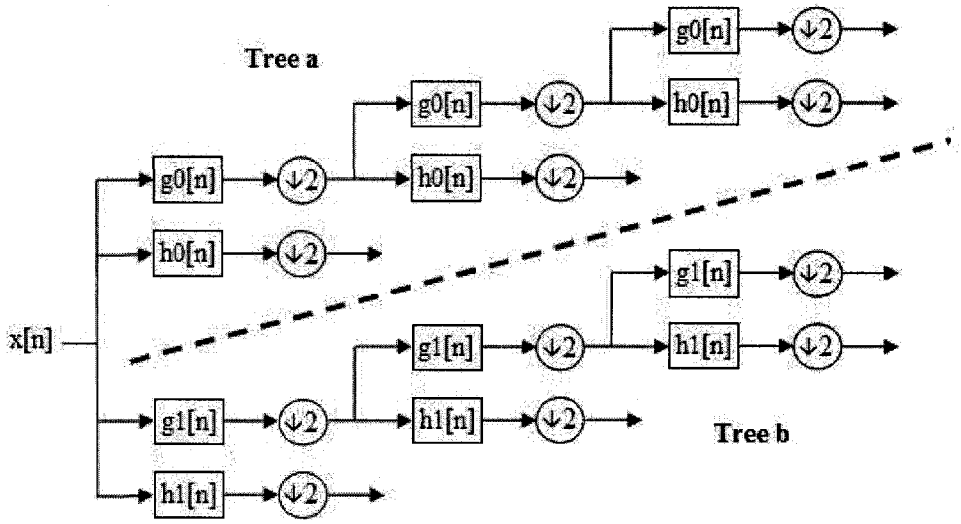

Figure 4.3: 2D DT-CWT filtering. $g[n]$ and $h[n]$: the lowpass and highpass filters which include complex filtering coefficients; index 0 and 1: tree a and tree b.

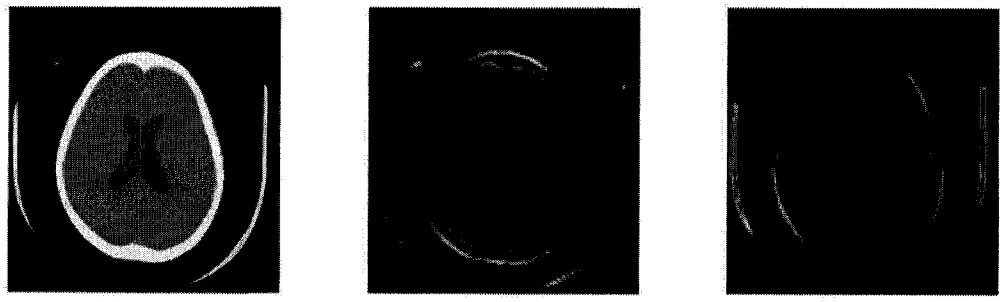

Figure 4.4: 2D DT-CWT example. From left to right (a-c): a) Original image; b) -75 subband at scale level 1; c) -15 subband at scale level 1

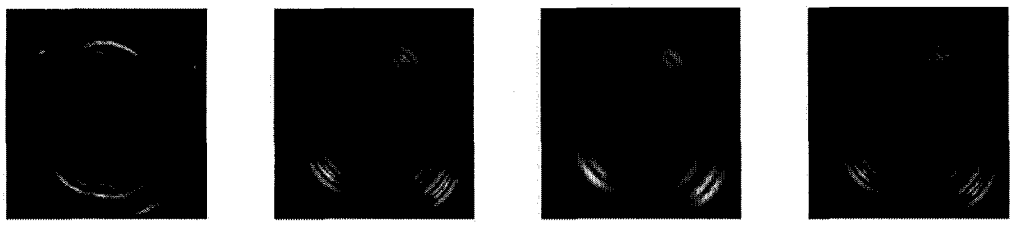

Figure 4.5: Problem using original DT-CWT coefficients. From left to right (a-d): a) Raw subband ($-75°$ at scale level 1); b) Gradient of (a); c) Normal 2D median filtered subband; d) Gradient of (c)

Directional Median Filtering

In this project, the features of $|D_{i,\theta}(x,y)|$ are not suitable for texture gradient calculation. According to [30], "the pitfall here though, is that the wavelet basis functions respond not only to extend patches of texture with a particular scale and orientation, but also to lines and step edges at that orientation across several scales". However, the features $|D_{i,\theta}(x,y)|$ can be used as the basis to extract *texture gradient* from the image. Figure 4.5 illustrates this problem, showing the response of the first level wavelet subband oriented at $-75°$.

After being passed to a normal 2D median filter which uses a 2D mask, double edges appear in the gradient magnitude image (Figure 4.5). If this image were to be passed to a watershed algorithm in its current form, a spurious narrow region would be grown along the boundary. Since the median filter is well known as a nonlinear edge-preserving smoothing or noise removal filter, directional median filtering is proposed by O'Callaghan [30] to solve the problem. The idea was developed because DT-CWT coefficients have six subbands at different orientations. A 2D normal me-

dian filter causes the smoothing effect on the subimage in all directions even though the subimage only includes detail information in a particular direction. Directional median filtering refers to adaptation to the orientation selected by each subband. Thus, both scale and orientation adaptation are employed in the median filtering.

The mathematical description of this median filtering on a particular subband with orientational angle θ at scale level i is introduced as:

$$S_{i,\theta}(x,y) = f_{MedFilt_\theta}(f_{MedFilt_{(\theta+\frac{\pi}{2})}}(|D_{i,\theta}(x,y)|)) \tag{4.7}$$

Equation (4.7) gives a clear procedure for applying two 1D median filterings on the subbands. The first median filter neighborhood extends in a line normal to the subband orientation. The second median filter is parallel to the subband orientation. The normal median filter is used to remove the step response (double edge effect) of the subbands and the parallel median filter is used to remove the noise of the subbands. Figure 4.6 shows the outcome of separable median filtering.

In practice, the size of the median filter is related to the extent of the filter bank impulse response at that level [30] and is chosen as $(7 + 2i)$, where i is the current level of the DT-CWT.

Interpolation

After directional median filtering, the new subbands $S_{i,\theta}(x,y)$ are now suitable for gradient extraction. Because of the nonlinear noise removal, the gradient of the new subbands no longer suffer from double edging around the intensity-step artifacts. The

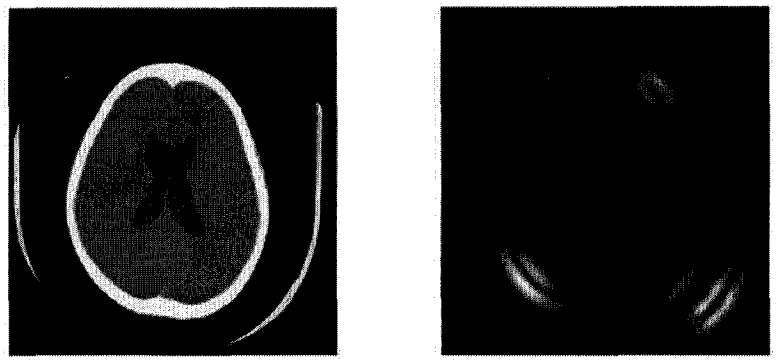

Figure 4.6: Separable median filtering. 1) The Original image. 2) The outcome after directional median filtering ($-75°$ subband at scale level 1)

Gaussian derivative function is applied as the gradient operator. Using the Gaussian derivative function not only estimates the gradient information of the image, but also mitigates the noise amplification. So, the texture gradient $G^{tex}_{i,\theta}(x,y)$ oriented at θ at scale level i of each subband is given by:

$$G^{tex}_{i,\theta}(x,y) = \sqrt{(S_{i,\theta}(x,y) * G'_x)^2 + (S_{i,\theta}(x,y) * G'_y)^2} \tag{4.8}$$

where G'_x and G'_y are the partial derivatives of the Gaussian in the x and y directions respectively and $*$ denotes convolution.

Since the input of the watershed algorithm is a single gradient map, all the different subbands need to be combined together to get a single texture-gradient function, $G^{tex}(x,y)$. According to [30], a simple weighted sum of magnitudes can be used to

do the combination. So:

$$G^{tex}(x,y) = \sum_{i,\theta} f_{interp}(\omega_{i,\theta} * \widehat{G}^{tex}_{i,\theta}(x,y)) \tag{4.9}$$

where

$$\widehat{G}^{tex}_{i,\theta}(x,y) = \frac{G^{tex}_{i,\theta}(x,y)}{\max_{x,y}(G^{tex}_{i,\theta}(x,y))} \tag{4.10}$$

$$\omega_{i,\theta} = \frac{N_i}{\sum_{x,y} \widehat{G}^{tex}_{i,\theta}(x,y)^2} \tag{4.11}$$

and N_i is the number of pixels in the subband image at level i. In Equation 4.9, f_{interp} is the simple linear interpolation function, zero insertion, since different subbands at different scale levels have different image sizes. When adding all the subbands together, we have to interpolate them and make sure they all have the same size as the original image.

4.2.2 Modulated Gradient

The modulated gradient is calculated by applying the morphological erosion operation on the images. This erosion is a modified method which is based on the DT-CWT subbands after directional median filtering, $S_{i,\theta}(x,y)$. In [16], the *texture activity* was introduced to obtain the modulated gradient. The purpose is to suppress the intensity gradient in textured areas but leave it unmodified in smooth regions. The measure of texture activity is described by

$$f_{Activity}(x,y) = e^{R_{half}(\frac{E_{tex}(x,y)}{\alpha} - \beta)} \tag{4.12}$$

where $R_{half}(\zeta)$ is half-wave rectification to suppress negative exponents and given by:

$$R_{half}(\zeta) = \begin{cases} 0 & \text{when } \zeta < 0 \\ \zeta & \text{when } \zeta \geq 0. \end{cases} \tag{4.13}$$

According to Equation (4.12), the activity measure is determined by two predefined parameters: α and β. O'Callaghan's [30] results showed that $\alpha = 2$ and $\beta = 7$ should be applicable to any 8-bit grayscale image. The texture energy, E_{tex}, is computed from the up-sampled subband features which are related to $S_{i,\theta}(x, y)$. So

$$E_{tex} = \sum_{i,\theta} f_{interp}(\epsilon_B(\frac{S_{i,\theta}(x, y)}{2^i})) \tag{4.14}$$

where, f_{interp} is the same interpolation function as in Equation 4.9, $S_{i,\theta}$ refers to Equation 4.7, and ϵ_B is the morphological erosion with B as the set of pixels in the structuring element. In this case the structuring element, B, is a square neighborhood of nine pixels, or the smallest approximately isotropic neighborhood. The erosion defined on grayscale images is:

$$[\epsilon_B(f)](x) = \min_{b \in B} f(x + b). \tag{4.15}$$

4.2.3 Modified Gradient for Watershed

Now, texture gradient and modulated gradient are combined to obtain a final gradient, $G_{watershed}(x, y)$, capturing all perceptual edges in the image. It is given by the

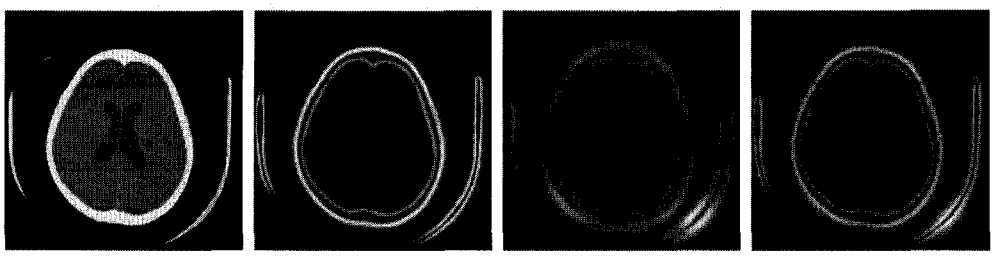

Figure 4.7: Modified gradient. From left to right (a to d): a) Original image; b) Modulated gradient; c) Texture gradient; d) Modified gradient

following equation:

$$G_{watershed}(x,y) = \frac{|\bigtriangledown G(x,y)|}{f_{Activity}(x,y) \times \omega_I} + \frac{G^{tex}(x,y)}{\omega_T} \tag{4.16}$$

where ω_T is the median value of the texture gradient, while ω_I is defined to be four times the median intensity gradient and $\bigtriangledown G(x,y)$ is the gradient of the original image. Figure 4.7 gives a good illustration of this process.

As mentioned in Chapter 3, the main drawback of the watershed algorithm is over-segmentation. Although the over-segmentation problem still exists even when the modified gradient is used, the result is much better than with the normal watershed algorithm. A post-processing technique called local minima suppression could be used to improve the segmentation results. The H-minima transform is used in this thesis. It modifies the gradient surface, suppressing shallow minima. Fig 4.8 compares the watershed segmentation computed with and without this modification.

This concludes the discussion of the first stage of the adaptive segmentation algorithm. Its flow chart is illustrated in Fig 4.9.

Figure 4.8: Local minima suppression. From left to right (a-b): a) without H-minima; b) with H-minima

4.2.4 Feature Extraction

All the methods in the previous section are gradient modifications and they are only a partial solution to the watershed over-segmentation problem in real medical images. There are inevitably some regions which should be merged together to form complete textures. A texture classification method is required to improve the performance of the segmentation algorithm.

A novel texture classification method is proposed in this thesis. Traditional texture classification is usually based on a single pixel value classification or fixed window-size classfication [13]. Both of these two methods treat a "small" area as the texture and try to extract the texture features. For example, in a single pixel value classification, one pixel is treated as a texture. The method uses the intensities of pixels to cluster the image into several different regions. Since a texture contains more than one gray level, different regions will overlap with each other. Fixed window-size classification performs better than the single pixel classification because a window is used to specify

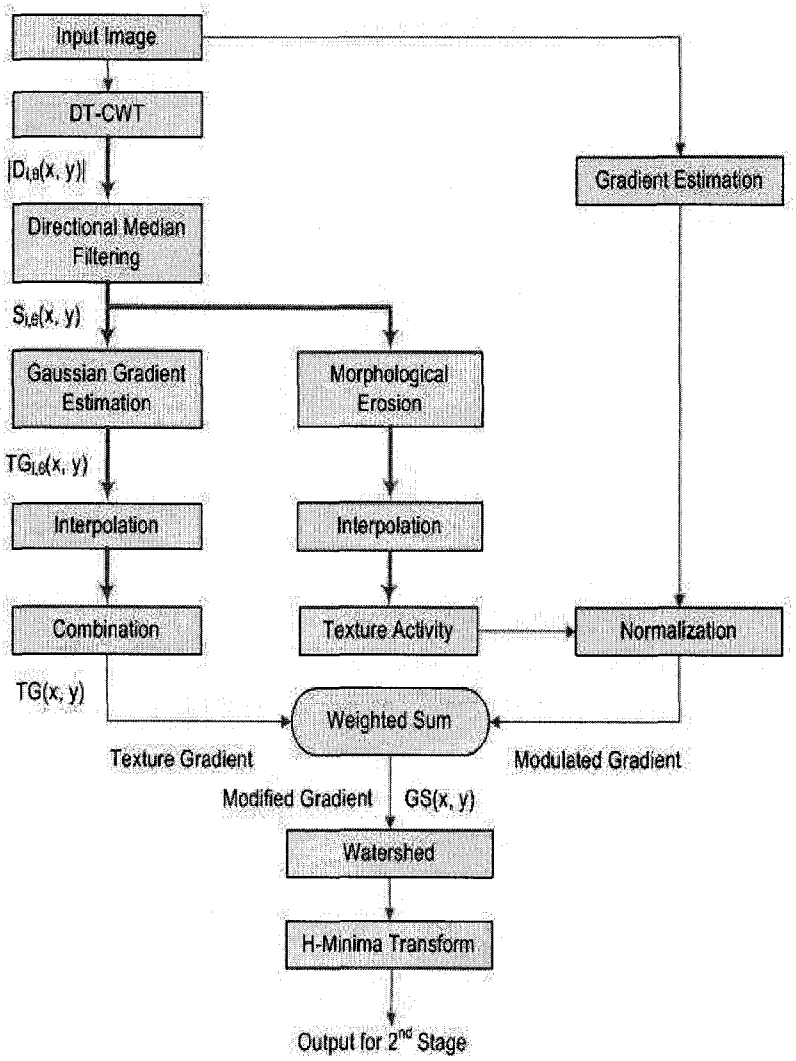

Figure 4.9: First stage of segmentation

a region which is used for feature extraction. When the window is inside of a texture, it could represent some characteristics of the current texture. But if the window crosses several textures, the features extracted from this window cannot accurately reflect the textures. Furthermore, the choice of the window size is an essential factor for the final classification results. If the window size is greater, the window might contain other textures and lead to inaccurate classification. If the window size is very small, the feature extraction method does not work very well since a very small sample size may not provide a good estimation of a feature of a texture. Therefore, a flexible window or region which includes sufficient and homogeous texture information should be used for classification. Because of the irregular boundaries of textures, it is usually very hard to find such a window or region. Fortunately, the adaptive segmentation algorithm provides a lot of regions which include sufficient and homogenous texture information. Since different textures have different characteristics, if the texture features are extracted from each region, different textures can be distinguished from each other based on the "distance" between them. Due to the approximate shift invariance and selective sensitivity of the DT-CWT, a new feature extraction method based on DT-CWT coefficients is proposed. This forms the second stage of the segmentation algorithm.

As discussed in Chapter 3, DT-CWT decomposes an image into seven subband images at each scale level. According to the theory of wavelet decomposition, only one subband image, which is filtered by the lowpass filter, is the low frequency approximation of the image. The other six subbands of the image are the detail information,

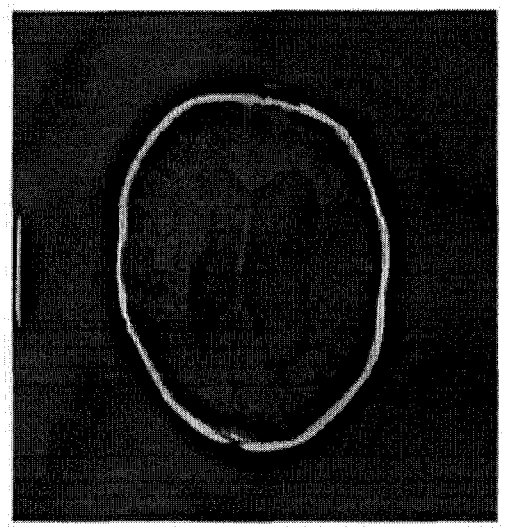

Figure 4.10: Reconstructed texture map. 4 scale level reconstruction with image size 512×512

and include a lot of texture information. For example, for scale level 4, one approximation subband image and 24 detail subbands can be obtained. Unlike other complex wavelets, DT-CWT allows perfect reconstruction.

A black image (pixel values of zero) with the same size as the approximation subband image was substituted for the approximation subband. Then, the inverse DT-CWT was applied to reconstruct a new image. This new image includes most of the texture information, while the approximate information is eliminated. The result of this reconstruction is referred to as a texture map and an example is given in Figure 4.10.

After the reconstruction of the texture map, the original image and the texture map are used as the inputs to the second stage. Since the preliminary results seg-

mented the image into several homogeneous regions, the distances between adjacent regions were computed and a distance matrix was built. In order to estimate the similarity between two samples, the Kolmogorov-Smirnov test (KS-test) from Matlab was used. Passing the texture map and the original image to the KS-test, two similarity matrices were obtained: $S1_{ks}$ for the texture map and $S2_{ks}$ for the original image. Now, the final similarity map S_{ks} for the merge process is given by the equation:

$$S_{ks} = S2_{ks} \times e^{(S1_{ks}-1)} \tag{4.17}$$

This equation was chosen so that the original image information would have the dominant effect when measuring the difference between regions and the texture information will be an accessory factor. The merge process is simply based on the final similarity matrix S_{ks}. Two regions which have the maximum value in S_{ks} are merged at this step. After merging, the labels for each region are updated and the new segmented image is used as the input of the second stage.

The second stage contains a loop as shown in Figure 4.11. First, the similarity map is calculated; then, the closest regions are merged; finally, the segmentation image is updated, and so on. A termination criterion is required to make the loop stop at the "best" segmentation step. The termination criterion used in the thesis was simple. When the maximum value in S_{ks} equals the minimum value in S_{ks}, it means that there are no two regions which are close and the program should stop. It was determined empirically that this termination criterion was suitable for most of the images tested. The process was refined to allow the physician to have some user

interaction to identify the ventricles if required. The whole flow chart of the second stage is shown in Figure 4.11.

To summarize, an image was roughly segmented at the first stage, then a texture classification stage was applied to optimize the outcome of the segmentation until a termination criterion was achieved. One example of the final segmentation results is shown in Figure 4.12. This figure also gives a comparison between the standard watershed and the adaptive watershed segmentation results.

4.3 Volume calculation

The volume calculation algorithm is applied after segmenting the ventricles from the CT images as detailed in Section 3.5.

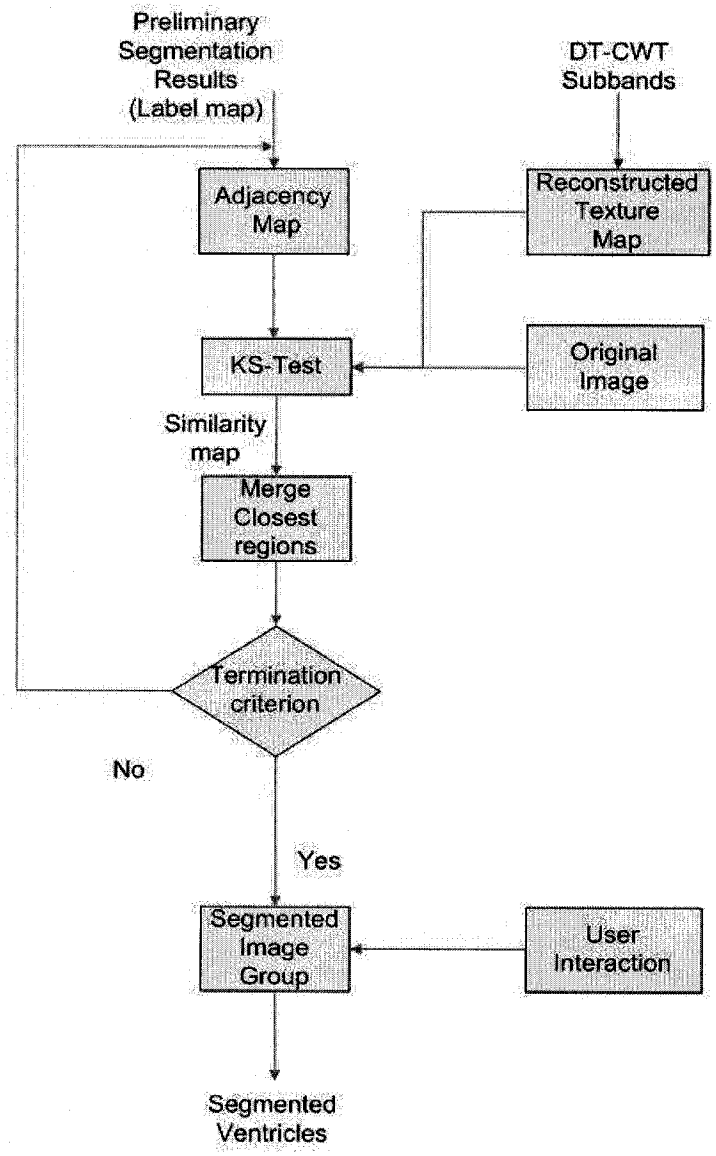

Figure 4.11: Second stage of segmentation

Figure 4.12: Final results. From left to right (a-b): a) Final segmentation result of the standard watershed algorithm; b) Final result of our adaptive watershed algorithm.

Chapter 5

Experimental Results

The algorithms were tested on both physical phantoms and clinical data. The following sections summarize and discuss the results. As mentioned in Chapter 4, the clinical diagnosis was known for all the test cases. In the following tables summarizing the results, the diagnostic comments refer to the following:

no hy: the patient was diagnosed as healthy.

hydrocephalus: the patient was diagnosed with hydrocephalus.

hy:stable: the patient has hydrocephalus but the hydrocephalus was stable between the two different scans.

hy:treated: the patient was diagnosed with hydrocephalus and was treated between scans.

5.1 Experimental Datasets

All the datasets were provided by the IWK Health Center in DICOM (Digital Imaging and Communication in Medicine) format. DICOM format consists of a group of specifications for messaging and communication between imaging machines. A complete DICOM file includes a header as well as the image data. The header stores all the information about the image data, for instance, the information about the patient, the technical parameters, and the image data. Three pieces of information important to the project were extracted from the header: the field of view (FOV), the slice thickness, and the image size. These values are used when calculating the volume of the ventricular system.

DICOM is the standard image format used in the medical system. The image data in DICOM format were burned onto a CD disk by a technician at the IWK. For confidentiality reasons, all the information about the patient was removed. Then, the anonymized data was used by the Department of Mathematics and Computing Science at Saint Mary's University where the image processing, algorithm design and algorithm validation were carried out. All algorithm development for this project was done in CVLAB [15], which is an image processing toolkit under Linux, and MATLAB. Since CVLAB does not support the DICOM format image, DICOM reader and viewer functions were developed for CVLAB. All the code was written either in the C language or using *.m files in Matlab.

The clinical data sets provided by the IWK Health Center were from cases that had previously been diagnosed and whose outcomes were known. The cases were

selected by a radiologist to reflect a range of likely clinical scenerios. The image data included 17 clinical cases and 5 cases from physical phantom models. All the image data was used to validate the algorithms developed for this project. Of the 17 clinical cases, eight cases were patients that had only one CT head scan carried out. Eight cases were patients that had CT scans of their head taken on at least two different occasions. One case had multiple scans. Each scan consisted of approximately 20 slices. For all the clinical cases, the hydrocephalus diagnosis was documented. For the eight cases with only one CT scan, a manual segmentation was also provided by the IWK. This manual segmentation was completed by a radiologist, Dr. Schmidt. This manual segmentation was used to validate the segmentation algorithm. The slice thickness for each case was identified and was 3mm, 5mm, or 7mm.

The five physical phantom models were built by Evans [9] at Dalhousie University, for the purpose of testing the volume calculation algorithm. The complex models consisted of a cylindrical agar "brain" and a complex fluid-filled space within it. The complex fluid-filled space was created from ice and its structure approached the complexity of the structure of the human ventricular system [39]. The choice of the agar and ice/water combination was based on several facts [9]. The density of the agar approximates the density of the tissue in the human brain and the density of ice and water approach that of the cerebral spinal fluid. The actual volume of the ventricles in the phantoms were known quantities since the volume of the water used to make the ice ventricles was measured. The physical phantoms were scanned in the same CT scanner that was used to collect the clinical data and similar settings were

used. Each phantom was given a complete CT scan at each of four different scanning angles, 0, 5,10, and 15 degrees. The slice thickness was either 3mm or 7mm and was identified for each case.

5.2 Results for Registration

The registration algorithm was tested on eight clinical cases that included two scans of the same patient but taken at different times. The clinical outcome of these cases was known. For three of the cases, the patients were diagnosed with hydrocephalus and the remaining cases were diagnosed as no hydrocephalus. The results of the registration algorithm are illustrated by the example given in Figure 5.1.

Figure 5.1: Registration results. From left to right (a to c): a) Reference image at time t_1; b) Floating image at time t_2; c) Registered image

In Figure 5.1, there is an obvious 3D movement between Figure 5.1a and Figure 5.1b. For example, the appearance of the eyeballs shows that there is a nodding movement and the whole skull has rotated. After registration, the eyeballs can be seen in the registered image and the position of the nose is corrected. Thus the 3D

displacement has been corrected. This is validation based on visual measurement.

A quantitative measure was also used to validate the registration algorithm. The measure is the Improvement Rate (IR). IR is a measure of the change in the difference between the initial image and subsequent image, giving an indication of how close the registered image is to the initial image, expressed as a percentage. The definition of IR is given in Equation (5.3).

$$D_{ref-float}(x,y) = |F_{ref}(x,y) - F_{float}(x,y)| \tag{5.1}$$

$$D_{ref-reg}(x,y) = |F_{ref}(x,y) - F_{reg}(x,y)| \tag{5.2}$$

$$IR = |\frac{\sum_{x,y} D_{ref-float}(x,y) - \sum_{x,y} D_{ref-reg}(x,y)}{\sum_{x,y} D_{ref-float}(x,y)}| \tag{5.3}$$

where F_{float} refers to the floating image at time, t_2, F_{ref} refers to the reference image at time t_1, and F_{reg} refers to the registered image. IR gives a more specific measure of how the registration algorithm aligns the images. Table 5.1 lists the IR results of the eight clinical cases.

Case name	IR	Diagnosis
p1	70.86%	no hy
p2	20.20%	hydrocephalus
p3	64.18%	no hy
p4	47.23%	no hy
p8	63.81%	no hy
p5	62.49%	no hy
p6	63.29%	hy:stable
p7	55.72%	hy:treated
Average IR	55.97%	

Table 5.1: IR results of registration algorithm

As seen in Table 5.1, the average IR of all the cases is 55.97%. For the three abnormal cases, IR varies from 20.20% to 63.29%. Case $p2$ has an IR of only 20.20%. In this case there is significant skull deformation caused by the hydrocephalus. The registered image, although aligned, is still dissimilar from the initial image. However, a 20.20% improvement after registration was still obtained. The rigid body registration algorithm shows a significant improvement in all cases. Moreover, the registration algorithm extended the work of Sun's thesis [39]. Sun's method did not allow for compensation of the nodding component of the movement. However, this nodding movement correction was achieved by using the adaptive registration algorithm in this thesis.

5.3 Results for Segmentation

The segmentation algorithm was validated using the similarity index. The similarity index, S, is given by the following equation:

$$S = 2 * \frac{|A_1 \cap A_2|}{|A_1| + |A_2|} \tag{5.4}$$

where A_1 and A_2 are the number of ventricle pixels in the images segmented using adaptive segmentation and manual segmentation respectively. According to [21], a similarity index $S \geq 0.7$ indicates excellent agreement. The average similarity index

S_{avg} of each case is computed using the formula:

$$S_{avg} = \sum_{i=1}^{N} \frac{S_i}{N} \tag{5.5}$$

where S_i is the similarity index of slice i and N is the total number of slices.

A total of eight cases were tested and the similarity index results between the adaptive segmentation and manual segmentation are listed in Table 5.2. Of all of the cases tested, six of them had normal ventricles and other two cases were known to have been diagnosed with hydrocephalus. Each case includes twenty to twenty-three slices and the thickness of the slice varied from case to case. Some cases had a slice thickness of $3.00mm$ and some had a slice thickness of $7.00mm$. After segmentation, the segmented images were changed to binary images (background in black and ventricle in white) in order to apply the similarity index.

Case name	Similarity index	Diagnosis
ps1	76.80%	no hy
ps2	77.08%	no hy
ps3	72.00%	no hy
ps4	72.37%	no hy
ps5	72.42%	no hy
ps6	74.93%	hydrocephalus
ps7	80.20%	no hy
ps8	74.60%	hydrocephalus
Average similarity	75.06%	
Similarity std	2.86%	

Table 5.2: Similarity results between adaptive segmentation and manual segmentation

In Table 5.2, the similarity index of all the cases was above 70%, which indicates that an appropriate segmentation was achieved [21]. In Figure 5.2, each point repre-

sents the average similarity index for each case. The range of the average similarity index for each case was from 72.00% to 80.20% and the average similarity of all the cases is 75.06%. Moreover, the standard deviation of the similarity data is only 2.86% which means the segmentation algorithm is stable for different cases. It doesn't vary significantly from one case to another.

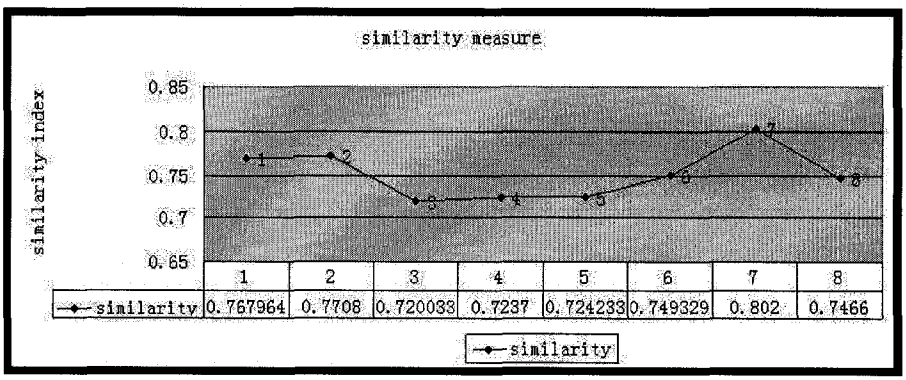

Figure 5.2: Graph of the similarity index results

5.4 Results for Volume Calculation

The purpose of this project is to develop a framework, using computer vision techniques, to help doctors diagnose hydrocephalus by measuring the volume of the ventricular system. The volume calculation algorithm, including the segmentation algorithm, was tested on physical phantoms to assess the algorithmic error. Then the complete framework including registration, segmentation, and volume calculation was tested on clinical cases.

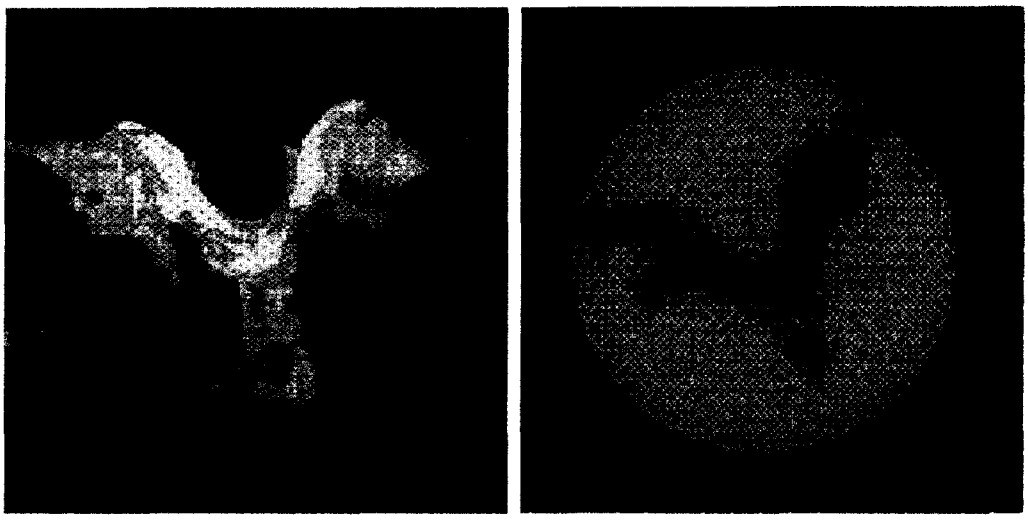

Figure 5.3: Physical phantoms. From left to right (a-b): a. phantom model; b. phantom slice

5.4.1 Physical Phantoms

The volume calculation algorithm [39] was tested on a set of physical models with measured volumes [9]. Five physical complex models (Fig 5.3) were used to validate the volume calculation algorithm including the segmentation algorithm. Each model, included both $3mm$ and $7mm$ slices in one case and was scanned at four different angles: 0 degrees, 5 degrees, 10 degrees, and 15 degrees. For each of these 5 complex models, Table 5.3 gives the known volume as well as the volume calculated by the algorithm. The first column is the actual volume of the 5 models; the second column gives the scanning angle in degrees; the third column lists the volume values computed by the volume algorithm including the adaptive segmentation algorithm; The fourth column documents the error as a percentage difference between the actual volume and the calculated volume.

Actual volume (cm^3)	Scanning angle in degrees	Calculated volume (cm^3)	Error (%)
102	0	105.27	+3.21%
	5	103.93	+1.89%
	10	104.85	+2.79%
	15	105.97	+3.90%
	average	105.00	+2.95%
	std	0.85	
112	0	111.19	-0.73%
	5	110.31	-1.51%
	10	106.70	-4.73%
	15	111.56	-0.39%
	average	109.94	-1.84%
	std	2.22	
130	0	132.73	+2.10%
	5	133.54	+2.73%
	10	134.79	+3.68%
	15	132.90	+2.23%
	average	133.49	+2.23%
	std	0.93	
101	0	99.97	-1.02%
	5	101.71	+0.70%
	10	103.09	+2.07%
	15	103.61	+2.59%
	average	102.09	+1.09%
	std	1.63	
88	0	86.10	-2.16%
	5	86.19	-2.06%
	10	89.43	+1.63%
	15	90.02	+2.29%
	average	87.94	+0.43%
	std	2.08	
Overall average error			0.97%

Table 5.3: Volume calculation results on physical phantoms (Actual volumes are from precision complex model cases [9])

In Table 5.3, the absolute values of all the errors are below 5% which is an acceptable range [40]. The signs in this table represent the direction of the estimation of the actual volume. A positive error means over-estimation of the actual volume. On the other hand, a negative error refers to the under-estimation of the actual volume. The overall average error is 0.97% which shows good estimation results between the volume calculation algorithm and the actual physical volume. The range of the standard deviations is from $0.85cm^3$ to $2.22cm^3$ which is reasonable compared with volumes of the complex models.

5.4.2 Clinical Cases

A total of 9 clinical cases with serial examinations which had been already diagnosed and whose follow up outcomes were also known, were used to validate the algorithm framework. The cases are divided into four groups: no hydrocephalus cases with *no hy* diagnosis comment, hydrocephalus cases with *hydrocephalus* comment, stable hydrocephalus cases with *hy:stable* comment, and treated hydrocephalus cases with the comment *hy:treated* in the last column. The three cases, $p9\_0, p9\_1, p9\_2$, are cases which follow the progress of a patient with more than two scans. For the cases diagnosed as no hydrocephalus, the ventricles were initially of regular size and did not exhibit any noticeable change in size between the two scans. The three hydrocephalus cases represent three typical situations. In case $p2$, the ventricle significantly increased its size during the time between scans. In case $p6$, the ventricular system is abnormally large but exhibits only a small change between the two scans. So, $p6$ is a stable

hydrocephalus case. In case $p7$, the patient had a seriously enlarged ventricle at scan time t_1. However, with careful treatment, the ventricle decreased significantly by the time of the second scan.

The cases labeled $p9\_0, p9\_1, p9\_2$ are cases which follow the progress of a patient with more than two scans and were taken over a period of two years. Initially, the patient had serious hydrocephalus. After 1 year, the patient got worse as indicated by the large increases in ventricle size. Then, the patient was treated successfully for hydrocephalus and the ventricle size decreased significantly. These $p9$ cases illustrate the the progress of the hydrocephalus.

Case name	Ventricle at tm1 (cm^3)	Ventricle at tm2 (cm^3)	Difference in volume (%)	Volume at tm2 after registration (cm^3)	Difference in volume after registration (%)	Diagnosis comments
p1	4.41	4.66	+5.62	4.25	-3.62	no hy
p2	71.68	169.83	+136.93	114.07	+59.14	hydrocephalus
p3	23.41	24.32	+3.91	23.85	+1.89	no hy
p4	4.42	5.63	+27.31	4.42	+0.05	no hy
p5	6.67	7.51	+12.63	6.65	-0.30	no hy
p6	29.78	30.11	+1.11	30.78	+3.37	hy:stable
p7	24.11	14.86	-38.35	17.43	-27.71	hy:treated
p8	10.61	12.61	+18.82	10.27	-3.16	no hy
p9_0	50.05	83.38	+66.60	70.63	+41.13	hydrocephalus
p9_1	83.38	76.90	-7.76	79.82	-4.27	hy:stable
p9_2	76.90	11.06	-85.62	15.94	-79.28	hy:treated

Table 5.4: Volume calculation results on clinical cases

Table 5.4 summarizes the data collected for each case. In this table, the first column is the name of each case; the second column represents the volume values of scan 1 which was taken at time t_1; the third column contains the volume values of scan 2 which was taken at time t_2; the forth column computes the difference in

Case name	Difference in volume (%)	Difference in volume after registration (%)	Diagnosis comments
p1	+5.62	-3.62	no hy
p3	+3.91	+1.89	no hy
p4	+27.31	+0.05	no hy
p5	+12.63	-0.30	no hy
p6	+1.11	+3.37	hy:stable
p8	+18.82	-3.16	no hy
p9_1	-7.76	-4.27	hy:stable
Average of magnitude	11.02	2.38	
Std of magnitude	9.29	1.67	

Table 5.5: Volume calculation results on clinical cases whose absolute values of the "Difference in volume after registration" are below 5%

Case name	Difference in volume (%)	Difference in volume after registration (%)	Diagnosis comments
p2	+136.93	+59.14	hydrocephalus
p7	-38.35	-27.71	hy:treated
p9_0	+66.60	+41.13	hydrocephalus
p9_2	-85.62	-79.28	hy:treated
Average of magnitude	81.88	51.82	
Std of magnitude	41.52	22.38	

Table 5.6: Volume calculation results on clinical cases whose absolute values of the "Difference in volume after registration" are above 5%

volume between time t_1 and time t_2; the fifth column contains the volume values of scan 2 after scan 2 was registered to scan 1. The difference in volume between time t_1 and time t_2 after registration is in the sixth column and the last column shows the diagnosic comments on each case. The formula to compute the difference in volume is defined

$$D_{volume} = \frac{v_{t_2} - v_{t_1}}{v_{t_1}} \tag{5.6}$$

where v_{t_1} and v_{t_2} are the volumes at time t_1 and time t_2, respectively. In Equation (5.6), D_{volume} is a signed number and might be positive or negative. The signs represent the direction of the volume change. For example, if D_{volume} is positive, it means the volume increases during the time. On the other hand, the volume decreases between two scans when D_{volume} is a negative number. In current clinical situations, doctors have the most difficulty diagnosing cases where the ventricular volume change in the range of 5%. Thus for this thesis, 5% was chosen as the dividing line between stable and diseased cases. Therefore, if the absolute value of the change is below 5%, there is no significant change of the volume.

From Table 5.4, some conclusions can be drawn:

1) For the *"no hy"* and *"hy:stable"* cases, the absolute values of some "differences in volume" are above 5% before the images are registered. According to the known diagnosis, the volume should not have changed significantly. However, the registration algorithm improves the results. After registration all the values for the no hydrocephalus or stable cases fall within the range ±5%.

2) For the no hydrocephalus cases, all the differences in volume after registration

vary from -4.27% to 3.37% and hence have absolute values that are below 5%. This means that there is no significant changes in the volume between scan 1 and scan 2. Thus the data output from the algorithm and the doctors' diagnosis are consistent.

3) For the three hydrocephalus cases, the algorithm output also correctly reflects the features of each case. For example, in case $p2$, the change in volume of $+59.14\%$ and the fact that the volume at time t_2 is larger than that at time t_1 indicates that this patient's ventricle size increased dramatically during the time between scans. In case $p6$, although the patient has hydrocephalus, the condition is stable and there is no significant change between scans. So, the $+3.37\%$ calculated change in volume probably reflects mainly algorithm accuracy. The patient did not get better and did not get worse. In case $p7$, the -27.71% change in volume again illustrates that there was a big change in the patient's condition between scan 1 and scan 2. The smaller volume value at t_2 and the sign of the number -27.71% indicate that the ventricle size decreased during the time period. This reflects the fact that the patient's condition improved between the two scans. This is consistent with the clinical knowledge that the patient's condition was treated.

4) For the special cases labeled $p9$, $p9\_0$ means the first stage of the process, $p9\_1$ is the middle stage of the process and $p9\_2$ is the final stage of the process. In $p9\_0$, with a high volume change of $+41.13\%$ and larger values at time t_2 ($70.63cm^3$), the algorithm predicts that the patient got worse from time t_1 to time t_2. At the middle stage ($p9\_1$), the ventricle size decreased slightly (-4.27%), but the patient still has serious hydrocephalus. The patient was diagnosed as stable during this time

period. At the final stage, a volume change of -79.28% to $15.94cm^3$ at time t_2 give a good indication that the patient is recovering from the hydrocephalus which is again consistent with the clinical diagnosis.

Table 5.4 gives a good summary of the clinical cases used to validate the framework in this thesis. To analyze the standard deviation of the results, the third and the last column of Table 5.4 are isolated and formed into two tables, Table 5.5 and Table 5.6. Table 5.5 has all the cases where the absolute value of the *"difference in volume after registration"* column is below 5% and Table 5.6 has all the cases where the absolute values are above 5%. Since we are interested in the magnitude of the change in volume, when the mean and the standard deviation were calculated for these two tables, absolute values were used. The standard deviation in Table 5.5 (1.67%) is much lower than that in Table 5.6 (22.38%). Therefore, for *no hy* and *hy:stable* cases, the algorithm is stable and varies rarely. On the other hand, for *hydrocephalus* and hy:treated cases, a much higher standard deviation is obtained which is consistent with the significant changes of the ventricular system during the process. Figure 5.4 graphically illustrates the results of all the clinical cases. In Figure 5.4, two clusters (all the blue dots) are clearly identified. One has a mean at 2.38 (the left red line) and the other has a mean at 51.82 (the right red line). For each red line, there are two green lines, one on each side of it. These two green lines give the range of the cluster values. For example, the cluster with the mean value 2.38 has the value range $[0.71, 4.05]$ and the other cluster with the mean value 51.82 has the value range $[29.44, 74.20]$. After considering these two ranges, 5% (shown as black line in Figure

5.4) was empirically selected as the dividing line between these two clusters.

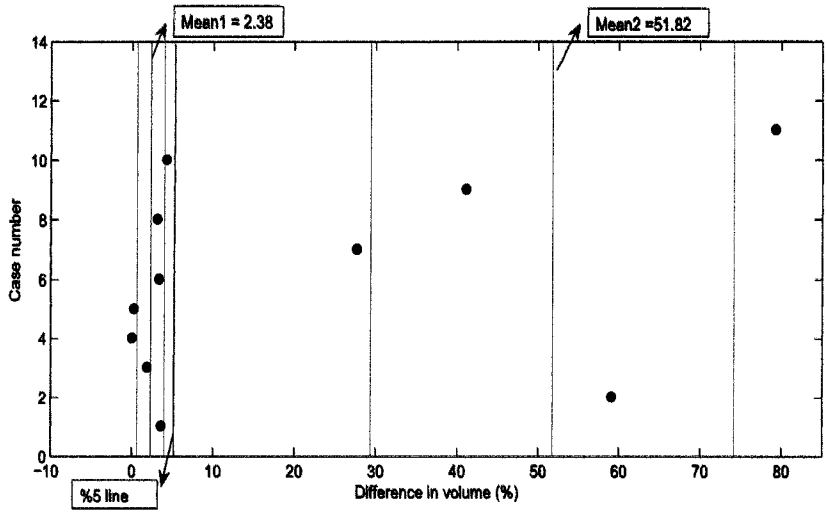

Figure 5.4: Graph of the clinical case results

In addition to the graph results of the clinical cases, the diagnostic performance of the set of clinical cases is discussed. In considering the diagnostic performance, there are four important concepts: true positive(TP), true negative(TN), false positive(FP), false negative(FN). In this thesis, TP refers to the number of cases which are diagnosed as hydrocephalus and the algorithm output also suggests a hydrocephalus diagnosis. TN is the number of the cases which are diagnosed as non-hydrocephalus and the algorithm also also suggests a non-hydrocephalus diagnosis. When the cases are non-hydrocephalus but the algorithm suggests a hydrocephalus diagnosis, it is FP. FN is the situation where the algorithm predicts a non-hydrocephalus diagnosis but the true diagnosis is hyrdrocephalus. According to these concepts, a diagnosis performance matrix is computed in Table 5.7. In Table 5.7, for both all the positive and negative examples, our algorithms give the correct predicted results.

	Predicted positive	Predicted negative	
Positive examples	6 (TP)	0 (FN)	6
Negative examples	0 (FP)	5 (TN)	5
	6	5	11

Table 5.7: Diagnostic performance analysis

Compared with existing similar systems, such as a visual measurement, Sze's estimation [40], and Sun's framework [39], the framework proposed in this thesis has its own robustness and efficiency. Compared to a visual measurement, the framework in this thesis uses a quantitative and more accurate method to measure the volume change of the ventricular system and thus reduces the chance of misdiagnosis caused by fatigue of the eyes.

Sze developed a volume calculation method but partial volume effects presented a serious drawback. In the method presented in this thesis, the image registration step reduces the partial volume effect and improves the accuracy of the results.

Sun's framework uses feature-based registration, region growing segmentation, and volume calculation. However, it only partially solved the displacement problem, and the region growing segmentation requires more user interactions than the proposed system. On the other hand, the framework in this thesis solved the displacement problem successfully without any user interactions. Moreover, the watershed-based segmentation algorithm provides acceptable segmentation results with minimal user interactions.

In conclusion, the algorithm predicted the doctor's diagnosis in all the cases that

were tested. Therefore, the adaptive registration, adaptive segmentation, and volume calculation algorithm in the framework are reasonable and the results of the project are acceptable.

Chapter 6

Conclusions and Future Work

In this thesis, a framework was implemented to measure the volume of the ventricular system to aid in the diagnosis of hydrocephalus. This framework consists of four important algorithms: wavelet analysis, a modified registration algorithm using a combination of the wavelet multiresolution pyramid and mutual information, an adaptive watershed segmentation with a novel feature extraction method based on the DT-CWT coefficients, and a volume calculation algorithm. In order to verify the assessment of the success of the algorithms used, an improvement rate (IR) was proposed to validate the registration algorithm and a similarity index used to test the segmentation algorithm. Finally, physical complex phantom models with known volumes and clinical cases with known diagnoses were used to validate the volume calculation algorithm.

The average IR of the no hydrocephalus cases is 55.97% which provides a good indication of how the registration algorithm compensated for both in-plane and out-

of-plane displacement between scans. The range of the similarity index for the 8 cases was 72.00% to 80.20% and the average similarity index of all the cases was 75.06% which was above 70.00% and fell into the acceptable range. This shows that the segmentation method worked well. For the volume calculation method on physical models, all the percent error rates were below 5% and the average percent error rate was 0.97%. These results are good since a 5% error rate was taken as acceptable [40]. For the volume calculation method on clinical models, the data reflect the correct diagnosis for each case, including the no hydrocephalus cases, hydrocephalus cases and special cases. Furthermore, graphical illustration and the diagnostic performance evaluation of all the clinical cases shows the success of our framework.

Some possible future work following from this thesis includes:

1. The adaptive segmentation algorithm does not have an accurate termination criterion. Future work could investigate the termination criterion.

2. The process used in this project segmented the ventricles in 2D first and then computed the volume in 3D. An alternate method would do the registration, segmentation, and volume calculation in 3D. As such, interpolation techniques between slices could be a potential topic for future research.

3. Although user interaction is minimized, the user still has to select the region of the ventricular system. Using the different textures found in the ventricles and other soft tissues, pattern recognition could be applied to help doctors find the correct textures quickly and accurately.

4. For the rigid body registration algorithm, 3D displacement was corrected in

this thesis. However, when there is a huge deformation in the ventricles in the scanned image, rigid body registration fails to find the most likely results. An alternative in this case is to automatically choose a landmark, such as the skull, then to optimize a region of interest that would be common to the two scans.

5. The diagnostic comments of all the test cases were known and documented. More cases with unknown diagnosis (or blind cases) need to be tested in the future to get an accurate measure of diagnosis performance.

6. Although our results have good potential for indicating the changes between two scans, more clinical data should be tested to validate the algorithms (registration, segmentation and volume calculation). In particular, more clinical data can be used to obtain the dividing line between stable and diseased cases using statistical methods. Also, clinical trails needs to be carried out.

7. A graphic user interface (GUI) could be implemented to facilitate the interaction of the users.

8. The registration algorithm runs quite slowly. Future work should include evaluation of more efficient implementations.

Bibliography

[1] Arivazhagan S, Ganesan L, 2003, *"Texture Classification Using Wavelet Transform"*, Pattern Recognition Letters 24 , pp.1513-1521.

[2] Ashtari M, Zito J.L, Gold B.I, Lieberman J.A, Borenstein M.T, Herman P.G, 1990, *" Computerized Volume Measurement of Brain Structure"*, Investigative Radiology, Vol.25, pp.789-805.

[3] Beucher S, 1994, *"Watershed, Hierarchical Segmentation and Waterfall Algorithm"*, in Mathematical Morphology and Its Applications to Image Processing. Boston, MA: Kluwer, pp.439-442.

[4] Beucher S, Meyer F, 1993, *"The Morphological Approach to Segmentation: the Watershed Transformation"*, in Mathematical Morphology and Its Applications to Image Processing, E.R. Dougherty, Ed. New York: Marcel Dekker, ch.12, pp.583-598

[5] Chang Y.L, Li X, 1994, *"Adaptive Image Region Growing"*, IEEE Transactions on Image Processing, Vol.3, No.6, pp.868-873.

[6] Cheriet M, Said J.N, Suen C.Y, 1998, *"A Recursive Thresholding Technique for Image Segmentation"*, IEEE Transactions on Image Prcoessing, Vol.7, No.6, pp.918-920.

[7] Collignon A, 1995, *"Automatic Multi-modality Medical Image Registration Based on Information Theory"*, Information Processing in Medical Imaging, pp.263-274.

[8] Dippel S, Stahl M, Wiemker R, Blaffert T, 2002, *"Multiscale Contrast Enhancement for Radiographies: Laplacian Pyramid Versus Fast Wavelet Transform"*, IEEE Transactions on Medical Imaging, Vol. 21, No. 4, pp.343-353

[9] Evans J, 2005, *"The Verfication of a Computer Algorithm Designed to Calculate the Volume of the Human Cerebral Ventricles Based on CT Images"*, Psychology Honours thesis, Dalhousie University.

[10] Farmer M.E, Jain A.K, 2005, *"A Wrapper-based Approach to Image Segmentation and Classification"*, IEEE Transactions on Image Processing, Vol.14, No.12, pp.2060-2072.

[11] Fei B, Wheaton A, Lee Z, Duerk J.L, Wilson D.L, 2002, *"Automatic MR Volume Registration and Its Evaluation for the Pelvis and Prostate"*, Physics in Medicine and Biology, Vol.47, No.5, pp.823-838.

[12] Frigui H, Krishnapuram R, 1999, *"A Robust Competitive Clustering Algorithm with Application in Computer Vision"*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.21, No.5, pp.450-465.

[13] Gonzalez R.C, Woods R.E, 2002, *Digital Image Processing*, 2nd Edition, Prentice Hall, pp.617-626.

[14] Grau V, Mewes A.U.J, Alcaniz M, Kikinis R, Warfield S.K, 2004, *"Improved Watershed Transform for Medical Image Segmentation Using Prior Information"*, IEEE Transactions on Medical Imaging, Vol.23, No.4, pp.447-458.

[15] Gregson P.H, 1996, *CVLAB Software*, Dalhousie University.

[16] Hill P.R, Canagarajah C.N, Bull D.R, 2003, *"Image Segmentation Using a Texture Gradient Based Watershed Transform"*, IEEE Transactions on Image Processing, Vol.12, No.12, pp.1618-1633.

[17] Jackway. P, 1996, *"Gradient Watersheds in Morphological Scale-space"*, IEEE Transactions on Image Processing, Vol.5, pp.913-921.

[18] Jung C.R, Scharcanski J, 2003, *"Adaptive Image Denoising and Edge Enhancement in Scale-space Using the Wavelet Transform"*, Pattern Recognition Letters 24, pp.965-971.

[19] Kak A.C, Slaney M, 1988, *Principles of Computerized Tomographic Imaging*, In Applied Mathematics, Society for Industrial and Applied Mathemtics, pp.263-274.

[20] Kaspersen J.H, Lango T, Lindset F, 2001, *"Wavelet-based Edge Detection in Ultrasound Images"*, Ultrasound in Med. and Bio., Vol. 27, No.1, pp.88-89.

[21] Kennedy D.N, Filipek P.A, Caviness V.S, 1997, *"Anatomic Segmentation and Volumetric Calculation in Nuclear Imagenetic Resonance Imaging"*, IEEE Transaction on Medical Imaging, Vol.8, No.1, pp.1-7.

[22] Kim B, Boes J.L, Bland P.H, Chenevert T.L, Meyer C.R, 1999, *"Motion Correctionn in fMRI via Registration of Individual Slices Into an Anatomical Volume"*, Magnetic Resonance in Medicine, Vol.41, pp.964-972.

[23] Kingsbury N.G, 2001, *"Complex wavelets for shift invariant analysis and filtering of signals"*, Journal of Applied and Computational Harmonic Analysis, Vol.10, No.3, pp.234-253.

[24] Kingsbury N.G, 1998, *"The Dual-tree Complex Wavelet Transform: a New Technique for Shift Invariance and Directional Filters"*, In Processing, 8th IEEE DSP Workshop, Bryce Canyon UT, USA, pp.86

[25] Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P, 1997, *"Multimodality Image Registration by Maximization of Mutual Information"*, IEEE Transactions on Medical Imaging, Vol.16, No.2, pp.187-198.

[26] Maes F, Vandermeulen D, Suetens P, 1999, *"Comparative Evaluation of Multiresolution Optimization Strategies for Multimodality Image Registration by Maximization of Mutual Information"*, Medical Image Analysis, Vol.3, No.4, pp.373-386.

[27] Maintz J.B.A, Viergever M.A, 1998, *"A Survey of Medical Image Registration"*, Medical Image Analysis, Vol.2, No.1, pp.1-36.

[28] Nelder J.A, Mead R, 1965, *"A Simplex Method for Function Minimization"*, Computer Journal 7, pp.308-313.

[29] Ng M.K, 2000, *"A Note on Constrained K-means Algorithm"*, Pattern Recognition, Vol.33, pp.515-519.

[30] O'Canllaghan R.J, Bull D.R, 2005, *"Combined Morphological-spectral Unsupervised Image Segmentation"*, IEEE Transactions on Image Processing, Vol.14, No.1, pp.49-62.

[31] Otsu N, 1978, *"A Threshold Selection Method From Grey Level Histograms"*, IEEE Transactions on Systems, Man nd Cybernetics, Vol.8, pp.62-66.

[32] Pluim J.P.W, Maintz J.B.A, Viergever M.A, 2000, *"Image Registration by Maximization of Combined Mutual Information and Gradient Information"*, IEEE Transactions on Medical Imaging, Vol.19, No.8, pp.809-814.

[33] Rashid-Farrorkhi F, Liu K.J.R, Berenstein C.A, Walnut D, 1997, *"Wavelet-based Multiresolution Local Tomography"*, IEEE Transactions on Image Processing, Vol. 6, No.10, pp.1412-1430.

[34] Saxena S.C, Kumar V, Hamde S.T, 2002, *"Feature Extraction From ECG Signals Using Wavelet Transforms for Disease Diagnostics"*, International Journal of System Science, Vol.33, No.13, pp.1073-1085.

[35] Schurr P.H, C.E. Polke, 1993, *"Hydrocephalus"*, Oxford Medical Publications, pp.1.

[36] Shafarenko L, Petrou M, Kittler J, 1997, *"Automatic Watershed Segmentation of Randomly Textured Color Images"*, IEEE Transactions on Image Processing, Vol.6, No.11, pp.1530-1543.

[37] Soille P, 1999, *"Morphological Image Analysis, Principles and Applications"*, Berl Germany: Springer-Verlag.

[38] Studholme C, Hill D.L.G, Hawkes D.J, 1999, *"An Overlap Invariant Entropy Measure of 3D Medical Image Alignment"*, IEEE Transactions on Medical Imaging, Vol.17, No.4, pp.586-595.

[39] Sun Z.Y, 2005, *"Using Computer Vision Techniques on CT Scans to Measure Changes in Ventricular Volume to Aid in the Diagnosis of Hydrocephalus"*, Master of Applied Science Thesis, Saint Mary's University, Halifax, Canada.

[40] Sze R.W, Ghioni V, Weinberger E, Seidel K.D, Ellenbogen R.G, 2003, *"Rapid Computed Tomography Technique to Measure Ventricular Volumes in the Child With Suspected Ventruculoperitoneal Shunt Failure I: Validation of Technique With a Dynamic Phantom"*, Journal of Computer Assisted Tomography, Vol.27, pp.663-667.

[41] Thevenaz P, Unser M, 1997, *"Spline Pyramids for Inter-modal Image Registration Using Mutual Information"*, Wavelet Application in Signal and Image Processing, WA: SPIE Press.

[42] Vincent L, Soille P, 1991, *"Watershed in Digital Spaces: an Efficient Algorithm Based on Immersion Simulations"*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.13, No.6, pp.683-598.

[43] Weisstein, Eric W. *Kolmogorov-Smirnov Test*, From MathWorld–A Wolfram Web Resource, http://mathworld.wolfram.com/Kolmogorov-SmirnovTest.html.

[44] Viola P, Wells III W.M, 1995, *"Alignment by Maximization of Mutual Information"*, International Conference on Computer Vision, pp.16-23.

[45] Yang G.Z, Hansell D.M, 1997, *"CT Image Enhancement with Wavelet Analysis for the Detection of Small Airways Disease"*, IEEE Transactions on Medical Imaging , Vol. 16, No.6, pp.953-961.

Saint Mary's University
Halifax, Nova Scotia
Canada B3H 3C3

Patrick Power Library
tel 902.420.5534
fax 902.420.5561
web www.stmarys.ca

Copyright Permission Notice

Documentation regarding permissions for the use of copyrighted material in this thesis is on file at the Saint Mary's University Archives. Contact us if you wish to view this documentation:

Saint Mary's University, Archives
Patrick Power Library
Halifax, NS
B3H 3C3

Email: archives@smu.ca
Phone: 902-420-5508
Fax: 902-420-5561